Mixed Finite Element Analysis of Lognormal Diffusion and Multilevel Monte Carlo Methods

I. G. Graham*, R. Scheichl*, and E. Ullmann*

June 26, 2015

Abstract

This work is motivated by the need to develop efficient tools for uncertainty quantification in subsurface flows associated with radioactive waste disposal studies. We consider single phase flow problems in random porous media described by correlated lognormal distributions. We are interested in the error introduced by a finite element discretisation of these problems. In contrast to several recent works on the analysis of standard nodal finite element discretisations, we consider here mass-conservative lowest order Raviart-Thomas mixed finite elements. This is very important since local mass conservation is highly desirable in realistic groundwater flow problems. Due to the limited spatial regularity and the lack of uniform ellipticity and boundedness of the operator the analysis is non-trivial in the presence of lognormal random fields. We establish finite element error bounds for Darcy velocity and pressure, as well as for a more accurate recovered pressure approximation. We then apply the error bounds to prove convergence of the multilevel Monte Carlo algorithm for estimating statistics of these quantities. Moreover, we prove convergence for a class of bounded, linear functionals of the Darcy velocity. An important special case is the approximation of the effective permeability in a 2D flow cell. We perform numerical experiments to confirm the convergence results.

Keywords: random porous media, fluid flow, lognormal random fields, mixed finite elements, multilevel Monte Carlo

Mathematics Subject Classification: 65N15, 65N30, 65C05, 60H35, 35R60

1 Introduction

Interest in the analysis, discretisation and postprocessing of partial differential equations (PDEs) with random coefficients has risen sharply over the past decade. These equations are used in computer simulations of physical processes with uncertain inputs in science, engineering and industry. The goal is to obtain quantitative estimates of the effect of input data uncertainties in order to reliably evaluate simulation results. Typical output quantities of interest are the expected value and higher moments of the solution or the probability of certain events which can be expressed as integrals over the input sample space.

In this paper we are concerned with the quantification of uncertainties in the simulation of subsurface flows. This plays an important role, for instance, in the safety assessment of proposed long-term radioactive waste repositories. To this end, we consider single phase fluid flow in a porous medium with random, correlated permeability. The flow is governed by Darcy's law with random permeability and conservation of mass. For simplicity, the pressure/flow boundary data is

*Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, UK. Email: I.G.Graham@bath.ac.uk, R.Scheichl@bath.ac.uk, E.Ullmann@bath.ac.uk.

assumed deterministic. Mathematically, this problem can be formulated as an elliptic PDE with a random diffusion coefficient. A standard discretisation for these equations with respect to the physical variables is by finite elements (FEs) where the solution to the PDE is approximated on meshes constructed in the spatial domain of interest. The well-posedness and subsequent finite element error analysis of elliptic PDEs with sufficiently regular, uniformly positive and bounded random coefficients has been established over the past decade and is classical [?, ?, ?, ?, ?].

Unfortunately, these results cannot be applied to practical subsurface flow situations where the permeability is modeled as a correlated, lognormal random field. Notably, this model ensures positivity of the permeability field and also accounts for the large variations of the permeability observed in real world data sets. However, lognormal diffusion coefficients are not bounded uniformly over all realisations and a non-standard approach has to be taken to prove existence, uniqueness, and regularity results for the PDE solution. Moreover, depending on the covariance function the trajectories of such permeability fields are often only Hölder continuous with exponent $t \leq 1$ and thus the regularity of the solution is limited. The analysis of the primal formulation of the lognormal flow problem can be found in [?, ?, ?, ?, ?, ?].

One important quantity of interest in radioactive waste disposal is the time it takes radionuclides in case of an accidental damage of the waste repository to travel from the (center of the) repository to the boundary of a well-defined safety zone. To be useful for realistic groundwater flow problems, it is desirable to use locally mass-conservative discretisation schemes. Discretisation schemes that do not have this property, such as standard Lagrange finite elements, lead to unphysical approximations of the Darcy velocity and of particle trajectories, even for very simple models of the particle transport. In the framework of finite element discretisations local mass-conservation can be achieved by mixed finite element methods [?].

Thus the goal of this paper is to extend the finite element error analysis in [?, ?] to lowest order Raviart-Thomas mixed elements [?]. In particular, we consider linear functionals of the (Darcy) velocity which enables us to analyse, e.g., the FE error for the effective permeability in a simple 2D flow cell. In addition, we study the properties of a piecewise linear, discontinuous recovered pressure approximation first introduced in [?]. This enjoys a faster convergence when the lognormal field has trajectories with Hölder exponent $t > 1/2$. For less smooth fields the standard piecewise constant pressure approximation converges with the same rate. This latter result is not surprising, but we could not find it in the classical literature. It relies on a duality argument similar to the ones used in [?, ?].

Mixed formulations of elliptic PDEs with lognormal coefficients are covered by the analysis (for the Brinkman problem) in [?]. However, in that work stabilised FEs are used and the analysis is carried out under strong regularity assumptions on the solution. Here, we make no such a priori assumption and instead deduce the regularity of the solution from the regularity of the diffusion coefficient. Moreover, we use lowest order Raviart-Thomas mixed finite elements together with piecewise constant pressure elements which do not require stabilisation. Mixed formulations of elliptic PDEs with random coefficients are also studied in [?], where an error analysis in the framework of stochastic Galerkin discretisations is carried out. However, there the random coefficient is assumed uniformly bounded which is not the case for lognormal coefficients.

The stochastic discretisation of PDEs with random coefficients often relies on Karhunen-Loève expansions of the random inputs [?]. This is not the case in Monte Carlo type methods. Karhunen-Loève expansions can be used to transform a random PDE into a parametric one. However, the typically short correlation length in subsurface flow problems results in very high-dimensional stochastic parameter spaces. Integration over this parameter space, the core task in uncertainty quantification, by spectral methods [?, ?, ?, ?], such as stochastic Galerkin or stochastic collocation methods, is at least up-to-now restricted to a few tens or hundreds of parameters. To date, algorithms for really high-dimensional integration instead use Monte Carlo (MC) based approaches

since, crucially, the convergence rates of these methods are dimension independent.

In subsurface flow situations as described above we have rough, highly variable, large contrast permeabilities. The standard MC estimator is in general very expensive in this context since it is necessary to generate large numbers of samples and to solve on very fine spatial meshes to obtain acceptable accuracies. The Multilevel Monte Carlo (MLMC) method overcomes this difficulty by computing approximations to output statistics on a hierarchy of meshes corresponding to the spatial discretisations and not only on a single mesh. MLMC was introduced by Heinrich [**?**] for the approximation of parameter-dependent integrals in high-dimensions. More recently it has been applied by Giles [**?**] in the context of stochastic differential equations and has since been used for the approximation of output statistics of PDEs with random coefficients, see [**?**, **?**, **?**, **?**, **?**].

Another option to analyse the MLMC convergence is by interpreting this method as a sparse tensor product approximation [**?**]. However, we follow the works [**?**, **?**] where the convergence and complexity of MLMC is analysed in terms of the root mean square error of the MLMC estimator. This approach is based on assumptions on the decay of the expected value of the FE error and of the variance of the difference of FE approximations on two consecutive grids, see [**?**, Theorem 1]. These assumptions have been proved for the model elliptic PDE with lognormal diffusion coefficients and standard Lagrange-type finite elements in [**?**] (see also [**?**, **?**] for further extensions). Here, we extend the error analysis and thus also the convergence analysis for MLMC to mass-conservative, mixed finite element schemes which are more suitable for subsurface flow applications. The results also form a crucial ingredient in the analysis of Quasi Monte Carlo (QMC) and multilevel QMC methods for lognormal diffusion problems in mixed form [**?**, **?**].

The rest of this paper is organised as follows. In Section 2 we present the lognormal diffusion model problem in mixed form and its discretisation by lowest order Raviart-Thomas elements for the Darcy velocity and piecewise constant elements for the hydrostatic pressure. The regularity of the solution to the mixed formulation is a basic ingredient for the FE error analysis and is studied in Section 2.2. Mixed FE error estimates for the Darcy velocity and for the pressure are derived in Section 3. Here, we also study two post-processing scenarios for Raviart-Thomas mixed methods of practial interest. We derive error estimates for a class of linear, bounded velocity functionals and for a piecewise linear pressure recovery process first proposed in [**?**]. In Section 4 we apply the error estimates in the complexity analysis of the multilevel Monte Carlo method applied to our model problem. Finally, in Section 5 we confirm the theoretical results by numerical experiments.

Notation In the rest of this paper we will use "$\lesssim \ldots$" to denote "$\leq C \ldots$" where the generic constant $C > 0$ is independent of the characteristic finite element mesh size, the approximated function and the random diffusion coefficient. Constants which depend on a specific realisation of the random coefficient will be stated explicitly. This is necessary since our analysis follows closely the idea in [**?**, **?**] where the finite element error analysis is first performed for a fixed realisation. Error estimates are then extended to the entire sample space using Hölder's inequality.

The FE error analysis of PDEs with deterministic coefficients is well established to date and standard textbook methods can be used. In contrast, the analysis of PDEs with random coefficients is non-trivial and requires considerable care. Since the input data of the PDE is random it can happen that the constants in the standard error estimates depend on a specific realisation or data set. Mathematically speaking, the constants are random variables. If we are interested in uniform error estimates on the entire sample space it is thus necessary to derive explicit expressions for these random variables.

2 Lognormal diffusion: mixed formulation and regularity

We study a coupled first-order system of PDEs with random coefficient on a bounded, Lipschitz polygonal/polyhedral domain $D \subset \mathbb{R}^d$, $d = 2, 3$, stated in mixed form:

$$a^{-1}(\omega, \vec{x})\vec{q}(\omega, \vec{x}) + \nabla u(\omega, \vec{x}) = \vec{g}(\omega, \vec{x}), \tag{2.1}$$

$$\operatorname{div} \vec{q}(\omega, \vec{x}) = f(\omega, \vec{x}) \qquad \text{in } D\ . \tag{2.2}$$

For the sake of a transparent presentation we assume deterministic Dirichlet boundary conditions $u(\omega, \vec{x}) = u_\Gamma(\vec{x})$ on the entire boundary ∂D of D. Note, however, that our analysis carries through without much further work also to Neumann and mixed boundary conditions, as well as to random boundary data. The test problem in Section 5 will in fact have a set of mixed boundary conditions. Now, given a probability space $(\Omega, \mathcal{A}, P)$ with sample space Ω and a probability measure P, we require that (2.1) – (2.2) as well as the boundary conditions are satisfied for all samples $\omega \in \Omega$ P-almost surely (P-a.s.). Note that the σ-algebra $\mathcal{A}$ associated with Ω is generated by the collection of random variables $\{a(\cdot, \vec{x})\colon \vec{x} \in D\}$.

In the context of steady-state flow in a porous medium, $\vec{q}$ is the Darcy velocity, u is the hydrostatic pressure and a is the permeability. The empirical relation between pressure and velocity (2.1) is known as Darcy's law and (2.2) is the law of conservation of mass. In this paper, the coefficient $a(\omega, \vec{x})$ is assumed to be a *lognormal* random field, i.e. $\log a(\omega, \vec{x})$ is Gaussian with a certain mean $\mu(\vec{x})$ and covariance function $C(\vec{x}, \vec{y}) := \mathbb{E}[(\log a(\cdot, \vec{x}) - \mu(\vec{x}))(\log a(\cdot, \vec{y}) - \mu(\vec{y}))]$.

For every sample $\omega \in \Omega$, we define

$$a_{\min}(\omega) := \min_{\vec{x} \in \overline{D}} a(\omega, \vec{x}) \quad \text{and} \quad a_{\max}(\omega) := \max_{\vec{x} \in \overline{D}} a(\omega, \vec{x}).$$

For these to be well-defined and to be able to use the regularity results proved in [**?**], we make certain assumptions on the coefficient a, the source terms $\vec{g}$ and f and the boundary data u_Γ. To this end, let $\mathcal{B}$ denote a Banach space with norm $\|\cdot\|_{\mathcal{B}}$. Let $L^p(\Omega, \mathcal{B})$ denote the space of $\mathcal{B}$-valued random variables with finite p^{th} moment (with respect to the probability measure P) of the $\mathcal{B}$-norm. For brevity we write $L^p(\Omega, \mathbb{R}) =: L^p(\Omega)$. The space of Hölder-continuous functions with exponent t is denoted by $C^t(\overline{D})$; $H^s(D)$ is the Sobolev space of (fractional) order $s \in \mathbb{R}$ (see, e.g., [**?**, Chapter 7]); $H(\operatorname{div}, D)$ is the subspace of functions $\vec{v} \in L^2(D)^d$ where $\operatorname{div} \vec{v} \in L^2(D)$ (see, e.g., [**?**]) with norm denoted by $\|\cdot\|_{H(\operatorname{div})}$. For fixed $\omega \in \Omega$, to simplify the presentation, we will write v_ω instead of $v(\omega, \cdot)$ for any function v on $\Omega \times D$ and likewise $\vec{v}_\omega$ instead of $\vec{v}(\omega, \cdot)$. With these definitions we can now state our basic assumptions:

A1. $a_{\min} > 0$ P-a.s. and $1/a_{\min} \in L^p(\Omega)$, for all $p \in [1, \infty)$,

A2. $a \in L^p(\Omega, C^t(\overline{D}))$, for some $0 < t \leq 1$ and all $p \in [1, \infty)$,

A3. $u_\Gamma \in H^{1/2+t}(\partial D)$, $\vec{g} \in L^r(\Omega, H^t(D)^d)$, $f \in L^r(\Omega, H^t(D))$ for some $0 < t \leq 1$ and $r \in [1, \infty)$.

Note A2 implies that $a_{\max}(\omega)$ is well-defined and $a_{\max}(\omega) \in L^p(\Omega)$, for all $p \in [1, \infty)$.

Assumptions A1–A2 are satisfied for any lognormal diffusion coefficient a where the underlying Gaussian random field $\log(a)$ has a Lipschitz continuous, isotropic covariance function and a mean function that belongs to $C^t(\overline{D})$ (cf. [**?**, Proposition 2.4]). Moreover, $1/a_{\min} \in L^p(\Omega)$ for all $p \in [1, \infty)$ is proved in [**?**, Proposition 2.3]. The Hölder-regularity of the trajectories of $\log(a)$ (and thus of a) follows from the smoothness of its covariance function. For the exponential covariance function it can only be shown that trajectories $a_\omega \in C^t(\overline{D})$, for all $t < 1/2$ [**?**, § 2.3]. For Matérn covariances with $\nu \in (1/2, 1)$ we have $a_\omega \in C^t(\overline{D})$ for all $t < \nu$ [**?**, § 2.2]. For $\nu > 1$, the trajectories $a_\omega \in C^1(\overline{D})$ and thus A2 holds for $t = 1$ [**?**, Remark 4].

Remark 2.1 The assumption $f \in L^r(\Omega, H^t(D))$ in A3 can be generalised to $f \in L^r(\Omega, H^{\tilde{t}}(D))$ for some $0 \leq \tilde{t} \leq 1$ with $\tilde{t} \neq t$. We choose not to do this to simplify the presentation.

2.1 Weak formulation

To further analyse and to eventually discretise (2.1)-(2.2), we put it in weak form (see, e.g. [**?**] for details). Let $\omega \in \Omega$ be fixed and set $\mathcal{V} = H(\mathrm{div}, D)$ and $\mathcal{W} = L^2(D)$. Introducing, for all $\vec{\eta}, \vec{v} \in \mathcal{V}$ and $w \in \mathcal{W}$, the bilinear forms

$$m_\omega(\vec{\eta}, \vec{v}) := (a_\omega^{-1}\vec{\eta}, \vec{v})_{L^2(D)}, \qquad b(\vec{v}, w) := -(\mathrm{div}\, \vec{v}, w)_{L^2(D)},$$

and the linear functionals

$$G_\omega(\vec{v}) := \big(\vec{g}_\omega, \vec{v}\big)_{L^2(D)} - \int_{\partial D} u_\Gamma\, \vec{v}\cdot\vec{\nu}\, ds, \qquad F_\omega(w) := -\big(f_\omega, w\big)_{L^2(D)},$$

the weak form of (2.1)-(2.2) is to find $(\vec{q}_\omega, u_\omega) \in \mathcal{V}\times\mathcal{W}$ such that P-a.s.

$$\left.\begin{array}{rcl} m_\omega(\vec{q}_\omega, \vec{v}) + b(\vec{v}, u_\omega) &=& G_\omega(\vec{v}) \quad \text{for all } \vec{v}\in\mathcal{V}, \\ b(\vec{q}_\omega, w) &=& F_\omega(w) \quad \text{for all } w\in\mathcal{W}. \end{array}\right\} \tag{2.3}$$

Existence and uniqueness results for problem (2.3) for fixed $\omega \in \Omega$ are classical (see, e.g., [**?**, Theorem 4.2.3]). They rely on certain continuity, coercivity and inf-sup stability conditions being satisfied. In particular, we need all the bilinear forms and linear functionals to be bounded. The following bounds follow immediately from the Cauchy-Schwarz inequality and a trace result:

$$\begin{aligned} &\|m_\omega\|_{\mathcal{L}(\mathcal{V},\mathcal{V}')} \le \frac{1}{a_{\min}(\omega)}, \qquad \|b\|_{\mathcal{L}(\mathcal{V},\mathcal{W}')} \le 1, \\ &\|F_\omega\|_{\mathcal{W}'} \le \|f_\omega\|_{L^2(D)} \quad \text{and} \quad \|G_\omega\|_{\mathcal{V}'} \le \|\vec{g}_\omega\|_{L^2(D)} + \|u_\Gamma\|_{H^{1/2}(\partial D)}. \end{aligned} \tag{2.4}$$

The inf-sup stability of b is also classical, i.e.

$$\sup_{\vec{v}\in\mathcal{V}} \frac{b(\vec{v}, w)}{\|\vec{v}\|_{H(\mathrm{div})}} \ge k_0 \|w\|_{L^2(D)}, \quad \text{for all } \ w\in\mathcal{W}, \tag{2.5}$$

with a constant $k_0 > 0$ that is independent of ω and only depends on the shape of the domain D. Finally, to establish the coercivity of m_ω let us introduce

$$Z = \{\vec{v}\in\mathcal{V} : b(\vec{v}, w) = 0 \quad \text{for all} \quad w\in\mathcal{W}\}.$$

This subspace of $\mathcal{V}$ is called $\ker B$ in [**?**]. The bilinear form $m_\omega(\cdot,\cdot)$ is coercive on Z, i.e.

$$m_\omega(\vec{v}, \vec{v}) \ge \frac{1}{a_{\max}(\omega)} \|\vec{v}\|^2_{H(\mathrm{div})}, \qquad \text{for all } \ \vec{v}\in Z. \tag{2.6}$$

Proposition 2.2 *Under the Assumptions A1–A3, the family of problems* (2.3) *has a unique solution* $(\vec{q}, u)$ *with* $\vec{q}\in L^p(\Omega,\mathcal{V})$ *and* $u\in L^p(\Omega,\mathcal{W})$, *for all* $1 \le p < r$.

Proof. For $\omega\in\Omega$ fixed there is a unique solution $(\vec{q}_\omega, u_\omega)\in\mathcal{V}\times\mathcal{W}$. This follows from the continuity conditions (2.4), together with the inf-sup stability (2.5) of b and the coercivity (2.6) of m_ω, since $0 < \frac{1}{a_{\max}(\omega)} \le \frac{1}{a_{\min}(\omega)} < \infty$ for almost all $\omega\in\Omega$. That $\vec{q}\in L^p(\Omega,\mathcal{V})$ and $u\in L^p(\Omega,\mathcal{W})$ follows from standard stability estimates of $\|\vec{q}_\omega\|_{\mathcal{V}}$ and $\|u_\omega\|_{\mathcal{W}}$ (see, e.g., [**?**, Theorem 4.2.3]) together with (2.4)–(2.6), Assumptions A1–A3, and the Hölder inequality. □

2.2 Regularity of the solution

To prove convergence of finite element approximations to the solution of (2.3) we need to study the regularity of its solution. Due to the following equivalence between primal and dual formulations of (2.1)-(2.2), we can use the regularity results proved in [**?**, **?**].

Lemma 2.3 *Suppose Assumptions A1-A3 hold and $(\vec{q}_\omega, u_\omega) \in \mathcal{V} \times \mathcal{W}$ is the unique solution of (2.3), P-a.s. for $\omega \in \Omega$. Then $u_\omega \in H^1_{u_\Gamma}(D) := \{v \in H^1(D) : v = u_\Gamma \text{ on } \partial D\}$ and*

$$\Big(a_\omega \nabla u_\omega, \nabla \phi\Big)_{L^2(D)} = \Big(\tilde{f}_\omega, \phi\Big)_{L^2(D)} \qquad \text{for all } \ \phi \in H^1_0(D)\ , \tag{2.7}$$

where $\tilde{f}_\omega := f_\omega - \operatorname{div}(a_\omega \vec{g}_\omega)$. Moreover, we have

$$\|u_\omega\|_{H^1(D)} \lesssim \frac{1}{a_{\min}(\omega)} \left(\|f_\omega\|_{H^{-1}(D)} + a_{\max}(\omega)\|\vec{g}_\omega\|_{L^2(D)}\right) + |u_\Gamma|_{H^{1/2}(\partial D)}\,. \tag{2.8}$$

Proof. Let us fix $\omega \in \Omega$. Choose $\vec{v} \in H(\operatorname{div}, D)$ with $v_i \in C_0^\infty(D)$ and $v_j = 0$, for $j \neq i$. Then the first equation in (2.3) gives

$$\int_D u_\omega \frac{\partial v_i}{\partial x_i}\,\mathrm{d}\vec{x} = \int_D (a_\omega^{-1}(q_\omega)_i - (g_\omega)_i) v_i\,\mathrm{d}\vec{x}\,.$$

This ensures the existence of weak first-order partial derivatives of u_ω, such that

$$\nabla u_\omega = -(a_\omega^{-1}\vec{q}_\omega - \vec{g}_\omega) \tag{2.9}$$

and $u_\omega \in H^1(D)$. Now, using the second equation in (2.3) and integrating by parts, for any $\phi \in H^1_0(D)$,

$$\big(f_\omega, \phi\big)_{L^2(D)} = \big(\operatorname{div}\vec{q}_\omega, \phi\big)_{L^2(D)} = -\big(\vec{q}_\omega, \nabla\phi\big)_{L^2(D)} = \big(a_\omega \nabla u_\omega, \nabla\phi\big)_{L^2(D)} + \big(\operatorname{div}(a_\omega \vec{g}_\omega), \phi\big)_{L^2(D)}\,.$$

Moreover, the first equation in (2.3) and Green's formula tell us that if $\vec{v} \in (C^\infty(D))^d$ then

$$\int_{\partial D} u_\omega \vec{v}\cdot\vec{\nu}\,ds \;=\; \int_D \nabla u_\omega \cdot \vec{v} \;+\; u_\omega \operatorname{div}\vec{v}\,\mathrm{d}\vec{x} \;=\; \int_D (\vec{g}_\omega - a_\omega^{-1}\vec{q}_\omega)\cdot\vec{v} \;+\; u_\omega \operatorname{div}\vec{v}\,\mathrm{d}\vec{x} \;=\; \int_{\partial D} u_\Gamma \vec{v}\cdot\vec{\nu}\,ds\,.$$

Hence, $u_\omega = u_\Gamma$ on all the smooth parts of ∂D. The extension to Lipschitz polygonal boundaries is classical [**?**]. In summary, $u_\omega \in H^1_{u_\Gamma}(D)$ and by (2.9) it satisfies (2.7). Since

$$\|\tilde{f}_\omega\|_{H^{-1}(D)} = \|f_\omega - \operatorname{div}(a_\omega\vec{g}_\omega)\|_{H^{-1}(D)} \le \|f_\omega\|_{H^{-1}(D)} + a_{\max}(\omega)\|\vec{g}_\omega\|_{L^2(D)}\,,$$

the bound on the H^1-norm of u_ω is a consequence of the Lax-Milgram Lemma. □

Since the solution of (2.3) is also the solution of the second-order problem (2.7), we can use the regularity results established in [**?**, Prop. 3.1] and [**?**, Thm. 2.1] to deduce the regularity of the solution of (2.3). We only state the result for $d = 2$. Similar results can also be proved for $d = 3$ and for coefficients that are only piecewise $C^t(\overline{D})$ (see [**?**] for details).

Theorem 2.4 *Let $D \subset \mathbb{R}^2$ be a polygon whose largest interior angle is $\theta_{\max} \in (0, 2\pi)$ and suppose Assumptions A1-A3 are satisfied for some $0 < t < 1$. Then $(\vec{q}_\omega, u_\omega) \in H^s(D)^d \times H^{1+s}(D)$, for all $0 < s < \min(t, \frac{\pi}{\theta_{\max}})$, P-a.s. in $\omega \in \Omega$, and the following bounds hold:*

$$\|u_\omega\|_{H^{1+s}(D)} \;\lesssim\; C_{\mathrm{reg}}(\omega) \quad \text{and} \quad \|\vec{q}_\omega\|_{H^s(D)} \;\lesssim\; \|a_\omega\|_{C^t(\overline{D})}\left[C_{\mathrm{reg}}(\omega) + \|\vec{g}_\omega\|_{H^s(D)}\right]\,, \tag{2.10}$$

where

$$C_{\mathrm{reg}}(\omega) := \frac{a_{\max}(\omega)\|a_\omega\|^2_{C^t(\overline{D})}}{a^4_{\min}(\omega)}\left[\|f_\omega\|_{L^2(D)} + \|a_\omega\|_{C^t(\overline{D})}\left(\|\vec{g}_\omega\|_{H^s(D)} + \|u_\Gamma\|_{H^{1/2+s}(\partial D)}\right)\right]$$

Moreover, $\operatorname{div}\vec{q}_\omega \in H^s(D)$ *and* $\|\operatorname{div}\vec{q}_\omega\|_{H^s(D)} = \|f_\omega\|_{H^s(D)}$ *for all* $0 \le s \le t$.

Proof. Let us fix $\omega \in \Omega$. Since u_ω is also a solution of the primal problem (2.7) (cf. Lemma 2.3), the regularity for u_ω and the bound on $\|u_\omega\|_{H^{1+s}(D)}$ follow immediately from [**?**, Thm. 2.1] provided $\tilde{f}_\omega \in H^{-1+s}(D)$. To show that $\operatorname{div}(a_\omega \vec{g}_\omega) \in H^{-1+s}(D)$ we can use [**?**, Lemma A.2], i.e. for any $\phi \in C^t(\overline{D})$ and $\psi \in H^s(D)$ with $0 < s < t < 1$, it follows that

$$\|\phi\psi\|_{H^s(D)} \lesssim \|\phi\|_{C^t(\overline{D})}\|\psi\|_{H^s(D)} . \tag{2.11}$$

If we apply this estimate with $\phi = a_\omega$ and $\psi = (g_\omega)_i$, $i = 1, \ldots, d$, we get

$$\|a_\omega \vec{g}_\omega\|_{H^s(D)} \lesssim \|a_\omega\|_{C^t(\overline{D})}\|\vec{g}_\omega\|_{H^s(D)} .$$

Since div is a linear and continuous operator from $H^s(\mathbb{R}^d)^d$ to $H^{-1+s}(\mathbb{R}^d)$ (cf. [**?**, Remark 6.3.14(b)]) it follows, as in the proof of [**?**, Thm. 2.1], by a localisation argument that

$$\|\operatorname{div}(a_\omega \vec{g}_\omega)\|_{H^{-1+s}(D)} \lesssim \|a_\omega\|_{C^t(\overline{D})}\|\vec{g}_\omega\|_{H^s(D)} .$$

To bound $\|\vec{q}_\omega\|_{H^s(D)}$ we use (2.9) and again (2.11) with $\phi = a_\omega$ and $\psi = \frac{\partial u_\omega}{\partial x_i} + (\vec{g}_\omega)_i$, $i = 1, \ldots, d$. We get

$$\|\vec{q}_\omega\|_{H^s(D)} \lesssim \|a_\omega\|_{C^t(\overline{D})}(\|u_\omega\|_{H^{1+s}(D)} + \|\vec{g}_\omega\|_{H^s(D)}) .$$

The fact that $\operatorname{div}\vec{q}_\omega \in H^s(D)$ and that its $H^s(D)$-norm is equal to that of f_ω for all $0 \le s \le t$ follows from the second equation in (2.3) which implies $\operatorname{div}\vec{q}_\omega \equiv_{L^2} f_\omega$. □

Remark 2.5 For convex domains $D \subset \mathbb{R}^d$ and for input random fields a that are sufficiently smooth such that $t = 1$, e.g. for the Matérn covariance with $\nu > 1$ or for the Gaussian covariance, it is proved in [**?**, Theorem 2.1] that in fact $u_\omega \in H^2(D)$ P-a.s. in $\omega \in \Omega$. Provided f is also sufficiently smooth, that is $f \in H^1(D)$, then the Darcy velocity $\vec{q}_\omega \in H^1(D)^d$ and $\operatorname{div}\vec{q}_\omega \in H^1(D)$. In that case, the theoretical results below yield optimal error estimates. We will not state that explicitly every time.

3 Mixed finite element discretisation and error estimates

The mixed finite element discretisation of (2.3), for any $\omega \in \Omega$, is obtained by choosing finite dimensional subspaces $\mathcal{V}_h \subset \mathcal{V}$ and $\mathcal{W}_h \subset \mathcal{W}$ and seeking $(\vec{q}_{h,\omega}, u_{h,\omega}) \in \mathcal{V}_h \times \mathcal{W}_h$ such that

$$\left.\begin{aligned} m_\omega(\vec{q}_{h,\omega}, \vec{v}_h) \;+\; b(\vec{v}_h, u_{h,\omega}) &= G(\vec{v}_h) \quad \text{for all } \vec{v}_h \in \mathcal{V}_h, \\ b(\vec{q}_{h,\omega}, w_h) &= F_\omega(w_h) \quad \text{for all } w_h \in \mathcal{W}_h . \end{aligned}\right\} \tag{3.1}$$

Here, for simplicity, we restrict our attention to the case when $\mathcal{V}_h$ is the lowest order Raviart-Thomas space on simplices [**?**].

Let $\mathcal{T}_h$ denote a family of triangulations (meshes) of D into conforming d-simplices $T \in \mathcal{T}_h$ (i.e. triangles for $d = 2$ and tetrahedra for $d = 3$). We assume that $\mathcal{T}_h$ has maximum mesh size $h := \max_{T\in\mathcal{T}_h} \operatorname{diam}(T)$ and is non-degenerate as $h \to 0$, i.e. $\operatorname{diam}(T)/\rho_T \le \gamma_0$ for all $T \in \mathcal{T}_h$, where ρ_T denotes the radius of the largest closed ball contained in $\overline{T}$ with a constant γ_0 independent of h. Let $\mathcal{E}$ denote the set of all faces of the simplices in $\mathcal{T}_h$, that is triangle edges ($d = 2$) or tetrahedron faces ($d = 3$), respectively, with pre-described unit normal $\vec{\nu}_E$. Let $\mathcal{E}_I$ and $\mathcal{E}_D$ denote the subsets of $\mathcal{E}$ consisting of interior faces $E \subset D$ and boundary faces $E \subset \partial D$. For each element $T \in \mathcal{T}_h$ we define the space of shape functions

$$RT_0(T) := \{\vec{v} : T \to \mathbb{R}^d \mid \vec{v}(\vec{x}) = \vec{\alpha} + \beta\vec{x},\ \vec{\alpha} \in \mathbb{R}^d,\ \beta \in \mathbb{R}\} , \tag{3.2}$$

as well as the global space of discontinuous, piecewise RT_0 finite element functions

$$RT_{-1}(\mathcal{T}_h) := \{\vec{v} \mid \vec{v}|_T \in RT_0(T)\ \forall T \in \mathcal{T}_h\}\ . \tag{3.3}$$

Finally, the lowest order Raviart-Thomas space $\mathcal{V}_h$ is defined as

$$\mathcal{V}_h := RT_0(\mathcal{T}_h) := \{\vec{v} \in RT_{-1}(\mathcal{T}_h) \mid \vec{v}\cdot\vec{\nu}_E|_E \text{ is continuous across } E,\ \forall E \in \mathcal{E}_I\}\ . \tag{3.4}$$

Because of the special form of (3.2) it is easily shown that for any $\vec{v} \in RT_0(T)$ the normal component $\vec{v}\cdot\vec{\nu}_E$ is constant on any face E of T. Moreover, $\vec{v}_h \in \mathcal{V}_h$ can be completely determined by specifying the constant value of $\vec{v}_h \cdot \vec{\nu}_E$ for each $E \in \mathcal{E}_I \cup \mathcal{E}_D$. In addition, we define $\mathcal{W}_h$ to be the space of piecewise constant functions on D with respect to the mesh $\mathcal{T}_h$. The pair $(\mathcal{V}_h, \mathcal{W}_h)$ enjoy the commuting diagram property [**?**, p.109] which implies in particular that $\operatorname{div}\mathcal{V}_h = \mathcal{W}_h$.

It follows directly from (2.5) that b is inf-sup stable with the same constant $k_0 > 0$ on the finite element space $\mathcal{V}_h \times \mathcal{W}_h$. Similarly, m_ω is coercive on

$$Z_h \ = \ \{\vec{v}_h \in \mathcal{V}_h : b(\vec{v}_h, w_h) = 0 \quad \text{for all} \quad w_h \in \mathcal{W}_h\}\ , \tag{3.5}$$

as well with the same coercivity constant $a_{\max}(\omega)^{-1}$ as in (2.6). It follows (analogously to Proposition 2.2) from standard results on (discretised) mixed variational problems that the discrete variational problem (3.1) has a unique solution (see, e.g., [**?**, Theorem 4.2.3]) with finite p^{th} moments up to a certain order.

Proposition 3.1 *Under the Assumptions A1–A3, the family of discrete problems* (3.1) *has a unique solution* $(\vec{q}_h, u_h)$ *with* $\vec{q}_h \in L^p(\Omega, \mathcal{V}_h)$ *and* $u_h \in L^p(\Omega, \mathcal{W}_h)$*, for all* $1 \le p < r$*.*

3.1 Velocity approximation

We fix a realisation $\omega \in \Omega$ and compare the solutions of (2.3) and (3.1). To this end we use the following standard results from the theory of mixed finite elements in [**?**, **?**, **?**].

Proposition 3.2 *We have* $Z_h \subset Z$*, as well as*

$$\begin{aligned}
\|\vec{q}_\omega - \vec{q}_{h,\omega}\|_{H(\operatorname{div})} &\le \left(1 + \frac{1}{k_0}\right)\left(1 + \frac{a_{\max}(\omega)}{a_{\min}(\omega)}\right) \inf_{\vec{v}_h \in \mathcal{V}_h} \|\vec{q}_\omega - \vec{v}_h\|_{H(\operatorname{div})} \quad \text{and} \\
\|\vec{q}_\omega - \vec{q}_{h,\omega}\|_{L^2(D)} &\le \left(1 + \frac{a_{\max}(\omega)}{a_{\min}(\omega)}\right) \inf_{\vec{v}_h \in \mathcal{V}_h} \|\vec{q}_\omega - \vec{v}_h\|_{L^2(D)}\ .
\end{aligned}$$

Now we extend the classical error estimates for lowest order Raviart-Thomas interpolation [**?**] to fractional order spaces. Surprisingly there does not seem to be a good source of a proof for this result, although the analogous result for hexahedral elements in 3D is given in [**?**, Lemma 3.3]. Let $\Pi_h : H^t(D)^d \cap H(\operatorname{div}, D) \to \mathcal{V}_h$ be the interpolation operator onto the lowest order Raviart-Thomas space, as defined, for example, in [**?**, § 2.5]. Note that for $0 < t \le 1$ the space $H^t(D)^d$, $d = 2, 3$, is continuously embedded in $L^r(D)^d$, for some $r > 2$ (see, e.g. [**?**, Chapter 7]). This is sufficient to ensure that Π_h is well-defined even when t approaches 0 (see, e.g. [**?**, § 2.5]).

Lemma 3.3 *Let* $0 < t \le 1$*. Then, for any* $\vec{v} \in H(\operatorname{div}, D) \cap H^t(D)^d$

$$\|\vec{v} - \Pi_h\vec{v}\|_{L^2(D)} \ \lesssim \ h^t|\vec{v}|_{H^t(D)} + h\|\operatorname{div}\vec{v}\|_{L^2(D)}\,. \tag{3.6}$$

Moreover, for any $\vec{v} \in H(\operatorname{div}, D)$ *with* $\operatorname{div}\vec{v} \in H^t(D)$*,*

$$\|\operatorname{div}(\vec{v} - \Pi_h\vec{v})\|_{L^2(D)} \ \lesssim \ h^t\,\|\operatorname{div}\vec{v}\|_{H^t(D)}\,. \tag{3.7}$$

Proof. Let $\hat{T}$ be the $d-$dimensional unit simplex and let $\hat{\Pi}$ denote the Raviart-Thomas interpolation operator of lowest order on $\hat{T}$. Moreover let $\hat{\vec{v}}$ be any function in $H(\mathrm{div}\,,\hat{T}) \cap H^t(\hat{T})^d$. Then, since $\hat{\Pi}$ preserves constants, we have, for any constant vector $\hat{\vec{p}}$,

$$\begin{aligned} \|\hat{\vec{v}} - \hat{\Pi}\hat{\vec{v}}\|_{L^2(\hat{T})} &= \|(\hat{\vec{v}} - \hat{\vec{p}}) - \hat{\Pi}(\hat{\vec{v}} - \hat{\vec{p}})\|_{L^2(\hat{T})} \\ &\leq \|\hat{\vec{v}} - \hat{\vec{p}}\|_{L^2(\hat{T})} + \|\hat{\Pi}(\hat{\vec{v}} - \hat{\vec{p}})\|_{L^2(\hat{T})} \,. \end{aligned} \tag{3.8}$$

Remembering that the degrees of freedom of $\hat{\Pi}\hat{\vec{v}}$ are the integrals of the normal component of $\hat{\vec{v}}$ on each of the faces of $\hat{T}$, we can argue as in Lemma 3.15 and equation (3.39) of [**?**] to obtain

$$\begin{aligned} \|\hat{\Pi}(\hat{\vec{v}} - \hat{\vec{p}})\|_{L^2(\hat{T})} &\lesssim \|\hat{\vec{v}} - \hat{\vec{p}}\|_{H^t(\hat{T})} + \|\mathrm{div}\,(\hat{\vec{v}} - \hat{\vec{p}})\|_{L^2(\hat{T})}, \\ &= \|\hat{\vec{v}} - \hat{\vec{p}}\|_{L_2(\hat{T})} + |\hat{\vec{v}}|_{H^t(\hat{T})} + \|\mathrm{div}\;\hat{\vec{v}}\|_{L^2(\hat{T})} \,, \end{aligned} \tag{3.9}$$

for any $t > 0$. Inserting this into (3.8) we conclude that

$$\|\hat{\vec{v}} - \hat{\Pi}\hat{\vec{v}}\|_{L^2(\hat{T})} \lesssim \|\hat{\vec{v}} - \hat{\vec{p}}\|_{L_2(\hat{T})} + |\hat{\vec{v}}|_{H^t(\hat{T})} + \|\mathrm{div}\;\hat{\vec{v}}\|_{L^2(\hat{T})} \,. \tag{3.10}$$

Now as $\hat{\vec{p}}$ is arbitrary we can use the Bramble-Hilbert Lemma in fractional order spaces (e.g. [**?**, Prop. 6.1]) to estimate the L_2 term on the right hand side of (3.10), obtaining, in the end,

$$\|\hat{\vec{v}} - \hat{\Pi}\hat{\vec{v}}\|_{L^2(\hat{T})} \lesssim |\hat{\vec{v}}|_{H^t(\hat{T})} + \|\mathrm{div}\;\hat{\vec{v}}\|_{L^2(\hat{T})} \,. \tag{3.11}$$

Now for any simplex $T \in \mathcal{T}_h$, take any function $\vec{v} \in H(\mathrm{div}\,,T) \cap H^t(T)^d$. Let $F_T : \hat{T} \to T$ be the usual affine map with (constant) Jacobian DF_T and set $J_T = \det(DF_T)$. The Piola transform of $\vec{v}$ is $\hat{\vec{v}} = J_T(DF_T)^{-1}(\vec{v} \circ F_T)$, and simple scaling arguments (see, e.g. [**?**]) show that $|\hat{\vec{v}}|_{H^t(\hat{T})} \sim h_T^{1/2+t}|\vec{v}|_{H^t(T)}$. Moreover $\hat{\Pi}\hat{\vec{v}} = \widehat{\Pi_T\vec{v}}$, where Π_T is the Raviart-Thomas interpolation operator on T. Using these results in (3.11) we obtain

$$\|\vec{v} - \Pi_T\vec{v}\|_{L^2(T)} \lesssim h_T^t|\vec{v}|_{H^t(T)} + h_T\|\mathrm{div}\;\vec{v}\|_{L^2(T)} \,.$$

Then, squaring and summing over all $T \in \mathcal{T}_h$ we obtain (3.6).

The second bound is simpler since $\mathrm{div}\,(\Pi_h\cdot) = P_h(\mathrm{div}\,\cdot)$, where P_h is the L^2-orthogonal projection $P_h : \mathcal{W} \to \mathcal{W}_h$ (see e.g. [**?**, Prop. 2.5.2]). Then we have

$$\|\mathrm{div}\,(\vec{v} - \Pi_h\vec{v})\|_{L^2(D)} = \|(I - P_h)\mathrm{div}\,\vec{v}\|_{L^2(D)} \,,$$

and (3.7) follows directly by standard polynomial approximation results in fractional order spaces (see, e.g. [**?**]). □

The next theorem now follows immediately by combining Proposition 3.2 and Lemma 3.3 with Theorem 2.4.

Theorem 3.4 *Let the assumptions of Theorem 2.4 hold. Then we have P-a.s. in $\omega \in \Omega$ and for all $0 < s < \min(t, \frac{\pi}{\theta_{\max}}) < 1$ that*

$$\|\vec{q}_\omega - \vec{q}_{h,\omega}\|_{L^2(D)} \lesssim C_q(\omega)h^s \quad \text{and} \quad \|\vec{q}_\omega - \vec{q}_{h,\omega}\|_{H(\mathrm{div}\,,D)} \lesssim \left(C_q(\omega) + \frac{a_{\max}(\omega)}{a_{\min}(\omega)}\|f_\omega\|_{H^t(D)}\right) h^s \tag{3.12}$$

where

$$C_q(\omega) := \frac{a_{\max}(\omega)\|a_\omega\|_{C^t(\overline{D})}}{a_{\min}(\omega)} \left(C_{\mathrm{reg}}(\omega) + \|\vec{g}_\omega\|_{H^t(D)}\right) \,. \tag{3.13}$$

Due to Assumptions A1–A3 the following corollary is a consequence of Hölder's inequality: Recall that A1–A2 imply $a_{\max}$, $1/a_{\min}$, and $\|a_\omega\|_{C^t(\overline{D})}$ are in $L^p(\Omega)$, for all $p \in [1,\infty)$. Thus, any product of these quantities with a finite number of factors is again in $L^p(\Omega)$, for all $p \in [1,\infty)$, due to Hölder's inequality. Applying Hölder's inequality again together with Assumption A3 we conclude that $C_{\rm reg} \in L^p(\Omega)$ and $C_q \in L^p(\Omega)$, for all $p < r$.

Corollary 3.5 *Let $t \in (0,1)$ and $r \in [1,\infty)$ be the parameters in Assumptions A1–A3. Then, under the assumptions of Theorem 2.4 and with $0 < s < \min(t, \frac{\pi}{\theta_{\max}}) < 1$ we have*

$$\|\vec{q} - \vec{q}_h\|_{L^p(\Omega,L^2(D))} \lesssim h^s \quad \textit{and} \quad \|\vec{q} - \vec{q}_h\|_{L^p(\Omega,H(\mathrm{div},D))} \lesssim h^s\,, \quad \textit{for all} \;\; p < r.$$

3.2 Pressure approximation

Recall that realisations a_ω of the diffusion coefficient are in general only Hölder continuous with exponent $0 < t \le 1$ (cf. Assumption A2). Standard arguments from [**?**] would lead only to a pressure approximation error (for each sample $\omega \in \Omega$) of order h^s, for all $s < t$. We will show that the approximation error is actually of order h. To prove this estimate we first prove an auxiliary result, based on a duality argument. This estimates the difference between the mixed finite element approximation and the L^2-orthogonal projection of the exact pressure in $\mathcal{W}_h$. This is similar to the arguments in [**?**, **?**] but surprisingly we have not been able to find the error bound in (3.20) in the literature.

Let P_h denote the L^2-orthogonal projection introduced in the proof of Lemma 3.3. Let $w \in \mathcal{W}$ and note that since $\mathrm{div}\,\mathcal{V}_h = \mathcal{W}_h$, we also have

$$b(\vec{v}_h, P_h w) = -(\mathrm{div}\,\vec{v}_h, P_h w)_{L^2(D)} = -(\mathrm{div}\,\vec{v}_h, w)_{L^2(D)} = b(\vec{v}_h, w), \qquad \text{for all } \vec{v}_h \in \mathcal{V}_h\,. \tag{3.14}$$

Before we move on, recall the following classical result (see e.g. [**?**])

$$\|w - P_h w\|_{L^2(D)} \lesssim h \|w\|_{H^1(D)} \qquad w \in H^1(D)\,. \tag{3.15}$$

Lemma 3.6 *Under the assumptions of Theorem 2.4, for P-a.s. $\omega \in \Omega$*

$$\|P_h u_\omega - u_{h,\omega}\|_{L^2(D)} \;\lesssim\; \underbrace{\frac{a_{\max}(\omega)}{a^2_{\min}(\omega)} \left(\frac{a_{\max}(\omega)\|a_\omega\|^3_{C^t(\overline{D})}}{a^4_{\min}(\omega)}\, C_q(\omega) \;+\; \|f_\omega\|_{H^t(D)} \right)}_{=:C_u(\omega)} h^{2s}\,. \tag{3.16}$$

Proof. Following the argument in [**?**, p. 432], consider the dual mixed problem to find $(\vec{z}, \phi) \in \mathcal{V} \times \mathcal{W}$, s.t.

$$\begin{aligned} m_\omega(\vec{v}, \vec{z}) \;+\; b(\vec{v}, \phi) \;&=\; 0\,, && \text{for all } \vec{v} \in \mathcal{V}, \\ b(\vec{z}, w) \;&=\; (P_h u - u_h, w)_{L^2(D)}\,, && \text{for all } w \in \mathcal{W}\,. \end{aligned} \tag{3.17}$$

Recall that $P_h u - u_h$ and the dual solution $(\vec{z}, \phi)$ depend on the sample $\omega \in \Omega$ but we shall omit this relation in what follows. In the associated discrete dual problem we shall use the subscript h and replace $\mathcal{V}$ by $\mathcal{V}_h$ and $\mathcal{W}$ by $\mathcal{W}_h$, respectively. In particular, we denote by $(\vec{z}_h, \phi_h) \in \mathcal{V}_h \times \mathcal{W}_h$ the mixed finite element solution of (3.17).

Due to (3.14), it follows from the second equation in the discrete version of (3.17) that

$$\|P_h u - u_h\|^2_{L^2(D)} = b(\vec{z}_h, P_h u - u_h) = b(\vec{z}_h, u - u_h) = m_\omega(\vec{q}_h - \vec{q}, \vec{z}_h) \tag{3.18}$$

where in the last step we used (2.3) and (3.1) with test function $\vec{v} = \vec{v}_h = \vec{z}_h$. Now, using the first equation in (3.17) with test function $\vec{v} = \vec{q}_h - \vec{q}$ we further deduce that

$$\|P_h u - u_h\|^2_{L^2(D)} = m_\omega(\vec{q}_h - \vec{q}, \vec{z}_h - \vec{z}) + b(\vec{q} - \vec{q}_h, \phi) = m_\omega(\vec{q}_h - \vec{q}, \vec{z}_h - \vec{z}) + b(\vec{q} - \vec{q}_h, \phi - P_h\phi) \ .$$

In the final step we have simply used the second equations in (2.3) and (3.1) with test function $w = w_h = P_h\phi$, respectively. A simple application of the Cauchy-Schwarz inequality leads to

$$\|P_h u - u_h\|^2_{L^2(D)} \lesssim \frac{1}{a_{\min}(\omega)} \|\vec{q} - \vec{q}_h\|_{L^2(D)} \|\vec{z} - \vec{z}_h\|_{L^2(D)} + \|\mathrm{div}\,(\vec{q} - \vec{q}_h)\|_{L^2(D)} \|\phi - P_h\phi\|_{L^2(D)} \,. \quad (3.19)$$

Since m_ω is symmetric, (3.17) is a special case of (2.3) with data $\vec{g}_{dual} \equiv \vec{0}$, $f_{dual} = u_h - P_h u \in L^2(D)$ and $u_{\Gamma,dual} \equiv 0$. This allows us to bound the error $\|\vec{z} - \vec{z}_h\|_{L^2(D)}$ using Theorem 3.4. The bound (3.12) (applied to $\vec{q} = \vec{z}$) gives

$$\|\vec{z} - \vec{z}_h\|_{L^2(D)} \lesssim C_q(\omega)\, h^s = \frac{a^2_{\max}(\omega) \|a_\omega\|^3_{C^t(\overline{D})}}{a^5_{\min}(\omega)} \|P_h u - u_h\|_{L^2(D)}\, h^s \ .$$

Moreover, it follows from (3.15) (applied to $w = \phi$) and (2.8) that

$$\|\phi - P_h\phi\|_{L^2(D)} \quad \lesssim \quad \|\phi\|_{H^1(D)}\, h \lesssim \frac{1}{a_{\min}(\omega)} \|P_h u - u_h\|_{L^2(D)}\, h \,.$$

Using these bounds together with Theorem 3.4 (applied to $\vec{q}$) in (3.19) and dividing the result by $\|P_h u - u_h\|_{L^2(D)}$ we obtain the final bound in (3.16).

□

Using (3.15) and Lemma 3.6 we can now establish the improved convergence order for the pressure error for $0 < t < 1$.

Theorem 3.7 *Let the assumptions of Theorem 2.4 hold. Then, for all* $0 < s < \min(t, \frac{\pi}{\theta_{\max}}) < 1$,

$$\|u_\omega - u_{h,\omega}\|_{L^2(D)} \ \lesssim \ C_u(\omega)\, h^{\min(2s,1)} \quad (3.20)$$

where C_u *is defined in* (3.16). *Moreover,* $\|u - u_h\|_{L^p(\Omega, L^2(D))} \lesssim h^{\min(2s,1)}$, *for all* $p < r$.

Proof. It follows from (3.15) applied to $w = u_\omega$ and from Theorem 2.4 that

$$\|u_\omega - P_h u_\omega\|_{L^2(D)} \lesssim \|u_\omega\|_{H^1(D)}\, h \le \|u_\omega\|_{H^{1+s}(D)}\, h \lesssim C_{\mathrm{reg}}(\omega)\, h \le C_u(\omega)\, h$$

with $C_u(\omega)$ as defined in (3.16). Combining this bound with Lemma 3.6 the bound (3.20) follows immediately via the triangle inequality. Due to Assumptions A1–A3 the bound on the moments is then again a consequence of Hölder's inequality. □

3.3 Linear velocity functionals

Let $\mathcal{M} : \mathcal{V} \to \mathbb{R}$ be a continuous linear functional of the Darcy velocity $\vec{q}$. An important example is the effective permeability, which in a rectangular flow cell reduces simply to

$$\mathcal{M}(\vec{q}) = \frac{1}{|D|} \int_D q_1 \,\mathrm{d}\vec{x} \quad (3.21)$$

(see Section 5 for details). Another example is the average normal flux through some part of the boundary. Our goal in this section is to estimate the FE approximation error $|\mathcal{M}(\vec{q}) - \mathcal{M}(\vec{q}_h)|$.

Let $\omega \in \Omega$ be again fixed. We will again use a duality argument and so we introduce the following auxiliary problem: Find $(\vec{z}_\omega, \phi_\omega) \in \mathcal{V} \times \mathcal{W}$, s.t.

$$\left.\begin{array}{rcll} m_\omega(\vec{v}, \vec{z}_\omega) \;+\; b(\vec{v}, \phi_\omega) & = & \mathcal{M}(\vec{v}) \;, & \text{for all } \vec{v} \in \mathcal{V}, \\ b(\vec{z}_\omega, w) & = & 0 \;, & \text{for all } w \in \mathcal{W} \;. \end{array}\right\} \tag{3.22}$$

We consider a specific class of linear functionals $\mathcal{M} \in \mathcal{V}'$. Let $\vec{v} \in H(\mathrm{div}\,, D)$, then by the Riesz Representation Theorem

$$\mathcal{M}(\vec{v}) = \int_D \vec{\psi} \cdot \vec{v}\,\mathrm{d}\vec{x} + \int_D \mathrm{div}\,\vec{\psi}\,\mathrm{div}\,\vec{v}\,\mathrm{d}\vec{x}$$

for some $\vec{\psi} \in H(\mathrm{div}\,, D)$.

A4. We assume that $\mathcal{M} \in \mathcal{V}'$ has a Riesz representor $\vec{\psi} \in H^t(D)^d$ with $\mathrm{div}\,\vec{\psi} \in H^{1+t}(D)$, for some $0 < t \leq 1$.

Note that under this assumption we have, due to Green's formula,

$$\int_D \mathrm{div}\,\vec{\psi}\,\mathrm{div}\,\vec{v}\,\mathrm{d}\vec{x} = \int_{\partial D} \vec{v} \cdot \vec{\nu}\,\mathrm{div}\,\vec{\psi} ds - \int_D \vec{v} \cdot \nabla(\mathrm{div}\,\vec{\psi})\,\mathrm{d}\vec{x}$$

and hence we can write

$$\mathcal{M}(\vec{v}) = \int_D \vec{v} \cdot (\vec{\psi} - \nabla(\mathrm{div}\,\vec{\psi}))\,\mathrm{d}\vec{x} + \int_{\partial D} \vec{v} \cdot \vec{\nu}\,\mathrm{div}\,\vec{\psi}\, ds \;. \tag{3.23}$$

For the functional in (3.21), i.e. the effective permeability in a rectangular flow cell, the Riesz representor is $\vec{\psi} = (1/|D|, 0)^\top$. For this choice of $\vec{\psi}$ Assumption A4 is clearly satisfied.

We now have the following result.

Theorem 3.8 *Let Assumption A4 and the assumptions of Theorem 2.4 hold. Then,*

$$\|\mathcal{M}(\vec{q}) - \mathcal{M}(\vec{q}_h)\|_{L^p(\Omega)} \lesssim h^{2s},$$

for all $0 < s < \min(t, \frac{\pi}{\theta_{\max}}) < 1$ *and all* $p < r$.

Proof. Omitting again the dependence on $\omega \in \Omega$, let $(\vec{z}, \phi) \in \mathcal{V} \times \mathcal{W}$ be the solution to the dual mixed problem (3.22), and let $(\vec{z}_h, \phi_h) \in \mathcal{V}_h \times \mathcal{W}_h$ be the corresponding mixed FE solution.

Note first that the second equation in the discrete version of (3.22) implies $\mathrm{div}\,\vec{z}_h \equiv 0$ on all of D. Thus, by subtracting the first equation in (2.3) from that in (3.1) with test functions $\vec{v} = \vec{v}_h = \vec{z}_h$, respectively, this also implies that

$$m_\omega(\vec{q} - \vec{q}_h, \vec{z}_h) = -b(\vec{z}_h, u - u_h) = 0\,. \tag{3.24}$$

Now, choosing $\vec{v} = \vec{q} - \vec{q}_h$ in the first equation of (3.22) and using (3.24), as well as the bilinearity of m_ω, we obtain

$$\mathcal{M}(\vec{q} - \vec{q}_h) = m_\omega(\vec{q} - \vec{q}_h, \vec{z}) + b(\vec{q} - \vec{q}_h, \phi) = m_\omega(\vec{q} - \vec{q}_h, \vec{z} - \vec{z}_h) + b(\vec{q} - \vec{q}_h, \phi - P_h\phi)\,.$$

In the last step we used (as in the proof of Lemma 3.6) that $b(\vec{q} - \vec{q}_h, P_h\phi) = 0$. Thus, by the Cauchy-Schwarz inequality we finally get

$$|\mathcal{M}(\vec{q} - \vec{q}_h)| \;\leq\; \frac{1}{a_{\min}(\omega)} \|\vec{q} - \vec{q}_h\|_{L^2(D)} \|\vec{z} - \vec{z}_h\|_{L^2(D)} + \|\mathrm{div}\,(\vec{q} - \vec{q}_h)\|_{L^2(D)} \|\phi - P_h\phi\|_{L^2(D)}\,. \tag{3.25}$$

Note that (3.22) is a special case of (2.3) with data $\vec{g} = \vec{\psi} - \nabla(\operatorname{div}\vec{\psi}) \in H^t(D)^d$ (thanks to Assumption A4), $f \equiv 0$ and $u_\Gamma = -\operatorname{div}\vec{\psi} \in H^{1/2+t}(D)$ (due to the standard Trace Theorem). It follows again as in the proof of Lemma 3.6, by applying (3.15) and Lemma 2.3, and Theorem 3.4 to the dual problem (3.22) that

$$\begin{aligned}
\|\vec{z} - \vec{z}_h\|_{L^2(D)} &\lesssim \frac{[\max(a_{\max}(\omega), \|a_\omega\|_{C^t(\overline{D})})]^6}{a_{\min}^5(\omega)} \left(\|\vec{\psi} - \nabla(\operatorname{div}\vec{\psi})\|_{H^s(D)} + \|\operatorname{div}\vec{\psi}\|_{H^{1/2+s}(\partial D)} \right) h^s \,, \\
\|\phi - P_h\phi\|_{L^2(D)} &\lesssim \left(\frac{a_{\max}(\omega)}{a_{\min}(\omega)} \|\vec{\psi} - \nabla\operatorname{div}\vec{\psi}\|_{L^2(D)} + |\operatorname{div}\vec{\psi}|_{H^{1/2}(\partial D)} \right) h \,.
\end{aligned}$$

Substituting these two bounds in (3.25) and using Theorem 3.4 we obtain

$$|\mathcal{M}(\vec{q}_\omega - \vec{q}_{h,\omega})| \lesssim C_{\mathcal{M}}(\omega) \left(\|\vec{\psi} - \nabla(\operatorname{div}\vec{\psi})\|_{H^t(D)} + \|\operatorname{div}\vec{\psi}\|_{H^{1/2+t}(\partial D)} \right) h^{2s} \,,$$

for some constant $C_{\mathcal{M}}(\omega)$, that (as a function of ω) is a random variable $C_{\mathcal{M}} \in L^q(\Omega)$, for all $q < r$, due to Assumptions A1–A3 and Hölder's inequality. The result then follows. □

Remark 3.9 The treatment of pressure functionals $\mathcal{L} = \mathcal{L}(u)$ is similar. The dual problem in that case has $G_\omega \equiv 0$ and $F_\omega \equiv \mathcal{L}$. For ideas on how to generalise to Fréchet differentiable, nonlinear functionals see [**?**, § 3.2], where this is explained in the standard finite element case.

3.4 Hybridisation and recovered pressure approximation

The Galerkin matrix associated with the mixed formulation (3.1) is indefinite. This is a potential problem for both direct and iterative solvers and was considered a major drawback of mixed methods. *Hybridisation* overcomes this problem by introducing (additional) Lagrange multipliers. After block elimination of pressures and velocities, the system for the multipliers is symmetric positive definite and much smaller than the saddle point system associated with (3.1). It can be solved efficiently by multigrid methods (see, e.g. [**?**]). Crucially, the mixed velocity and pressure approximation can then be obtained by local post-processing.

For lowest order Raviart-Thomas elements the additional Lagrange multipliers are piecewise constant along interelement boundaries in the finite element triangulation. They are used in conjunction with the space $RT_{-1}(\mathcal{T}_h)$ of discontinuous, piecewise RT_0 functions (w.r.t. to $\mathcal{T}_h$) for the velocity approximation defined in (3.3), to enforce the continuity of the normal component of the velocities across interelement boundaries in a weak sense. This approach was originally proposed in [**?**] as an efficient implementation technique for mixed methods in linear elasticity. However, in [**?**] it was proved that the Lagrange multipliers contain extra information and can be used to construct a more accurate piecewise linear recovered pressure approximation. In the remainder of this section, we analyse the FE error of the recovered pressure approximation. This is a novel application of this method in the case of random coefficients. In addition, the diffusion coefficient has limited regularity which requires working in non-integer order spaces.

More recently, the Lagrange multipliers have been characterised as the solution of a variational problem in which velocity and pressure do not appear, leading to a new approach to the error analysis of hybridised mixed methods which gives error estimates for the Lagrange multipliers without using error estimates for the other variables [**?**, **?**]. In our analysis here, however, we follow the traditional approach in [**?**]. We mention that in [**?**] only the case $D \subset \mathbb{R}^2$ and Raviart-Thomas elements of even orders k have been described in detail, but the construction and the analysis extend also to three space dimensions for lowest order elements ($k = 0$). The key is again to use the results in [**?**] for a fixed realisation $\omega \in \Omega$, keeping track of the precise dependence of the constants on ω.

Let $\omega \in \Omega$ be fixed and let $M_{-1}(\mathcal{E})$ denote the space of all piecewise constant functions on the skeleton $\mathcal{E} = \mathcal{E}_I \cup \mathcal{E}_D$. The space of all functions in $M_{-1}(\mathcal{E})$ that vanish on boundary faces $E \in \mathcal{E}_D$ is denoted $M_{-1}(\mathcal{E}_I)$. We modify the standard Raviart-Thomas mixed Galerkin equations (3.1) by using velocity functions $\underline{\vec{q}}_{h,\omega} \in RT_{-1}(\mathcal{T}_h)$ together with multipliers $\lambda_{h,\omega} \in M_{-1}(\mathcal{E}_I)$ and pressure functions $\underline{u}_{h,\omega} \in \mathcal{W}_h$. In addition we introduce the bilinear form $b_T(\vec{v}, w) := -(\operatorname{div} \vec{v}, w)_{L^2(T)}$. We then seek $(\underline{\vec{q}}_{h,\omega}, \underline{u}_{h,\omega}, \lambda_{h,\omega}) \in RT_{-1}(\mathcal{T}_h) \times \mathcal{W}_h \times M_{-1}(\mathcal{E}_I)$ such that

$$\left.\begin{aligned} m_\omega(\underline{\vec{q}}_{h,\omega}, \vec{v}_h) + \sum_{T\in\mathcal{T}_h} b_T(\vec{v}_h, \underline{u}_{h,\omega}) + \sum_{E\in\mathcal{E}_I} (\lambda_{h,\omega}, \vec{v}_h \cdot \vec{\nu}_E)_{L^2(E)} &= G_\omega(\vec{v}_h) \quad \forall\, \vec{v}_h \in RT_{-1}(\mathcal{T}_h), \\ \sum_{T\in\mathcal{T}_h} b_T(\underline{\vec{q}}_{h,\omega}, w_h) &= F_\omega(\vec{v}_h) \quad \forall\, w_h \in \mathcal{W}_h, \\ \sum_{E\in\mathcal{E}_I} (\mu_h, \underline{\vec{q}}_{h,\omega} \cdot \vec{\nu}_E)_{L^2(E)} &= 0 \qquad \forall\, \mu_h \in M_{-1}(\mathcal{E}_I)\ . \end{aligned}\right\} \tag{3.26}$$

By construction it is clear that the hybridised mixed system (3.26) has a unique solution $(\underline{\vec{q}}_{h,\omega}, \underline{u}_{h,\omega}, \lambda_{h,\omega})$. Moreover, we have $\underline{\vec{q}}_{h,\omega} = \vec{q}_{h,\omega}$ and $\underline{u}_{h,\omega} = u_{h,\omega}$ where $(\vec{q}_{h,\omega}, u_{h,\omega})$ denotes the solution of the standard Raviart-Thomas mixed Galerkin equations (3.1). Hence our analysis of the mixed formulation in Sections 3.1–3.3 carries over to the hybridised formulation and we drop the lower bars in (3.26) in the rest of this section. It only remains to analyse the convergence of the piecewise linear recovered pressure approximation, which we define below.

Before defining the recovered approximation, we derive bounds for the difference between $\lambda_{h,\omega}$ and the trace of u_ω on the interior element boundaries $\mathcal{E}_I$. To this end we define the following inner product

$$(\mu_h, \rho_h)_{0,h} := \sum_{E\in\mathcal{E}} (\mu_h, \rho_h)_{L^2(E)} \tag{3.27}$$

and the corresponding broken norm $|\mu_h|_{0,h} := (\mu_h, \mu_h)_{0,h}^{1/2}$ on $M_{-1}(\mathcal{E})$. Since $u_\omega \in H_0^1(D)$, we know that $u_\omega|_{\mathcal{E}} \in L^2(\mathcal{E})$ and we denote by $P_h^{\mathcal{E}}$ the orthogonal projection from $L^2(\mathcal{E})$ onto $M_{-1}(\mathcal{E})$ with respect to the inner product (3.27). We then have the following approximation result.

Lemma 3.10 *For every element $T \in \mathcal{T}_h$ and every edge E of T*

$$\|\lambda_{h,\omega} - P_h^{\mathcal{E}} u_\omega\|_{L^2(E)} \lesssim \frac{1}{a_{\min}(\omega)} h_T^{1/2} \|\vec{q}_\omega - \vec{q}_{h,\omega}\|_{L^2(T)} + h_T^{-1/2} \|u_{h,\omega} - P_h u_\omega\|_{L^2(T)}\,, \tag{3.28}$$

where $h_T := \operatorname{diam}(T)$ and P_h is the L^2-orthogonal projection from $\mathcal{W}$ to $\mathcal{W}_h$ defined in Section 3.2.

Proof. We follow the proof of [**?**, Theorem 1.4] and extend it to the stochastic case and to $d = 3$ space dimensions. Consider an element $T \in \mathcal{T}_h$ and a face $E \subset \partial T$. Let $(\vec{q}_{h,\omega}, u_{h,\omega}, \lambda_{h,\omega})$ denote the solution of (3.26) and $(\vec{q}_\omega, u_\omega)$ the solution of (2.3), respectively. We omit the dependence on ω in the proof.

Since a function $\vec{v}_h \in RT_0(T)$ is uniquely determined by the (constant) value $\vec{v}_h \cdot \vec{\nu}_E$ on $E_i \subset \partial T$, $i = 1, \ldots, d+1$, it is clear that there exists a unique $\underline{\vec{\delta}}_h \in RT_0(T)$ such that

$$\left.\begin{aligned} \underline{\vec{\delta}}_h \cdot \vec{\nu}_E &= \lambda_h - P_h^{\mathcal{E}} u \quad \text{on } E, \\ \underline{\vec{\delta}}_h \cdot \vec{\nu}_{E'} &= 0 \qquad\quad \text{on } E' \in \partial T \setminus E\ . \end{aligned}\right\} \tag{3.29}$$

Moreover, since $\underline{\vec{\delta}}_h \in RT_0(T)$, we have

$$\|\operatorname{div} \underline{\vec{\delta}}_h\|_{L^2(T)} \ + \ h_T^{-1} \|\underline{\vec{\delta}}_h\|_{L^2(T)} \ \lesssim \ h_T^{-\frac{1}{2}} \|\lambda_h - P_h^{\mathcal{E}} u\|_{L^2(E)}\,. \tag{3.30}$$

An elementary calculation shows that this bound holds on the reference element $\widehat{T}$. A simple scaling argument then gives (3.39). As in the proof of Lemma 2.3, a local application of Green's formula gives

$$\int_T (\vec{g} - a^{-1}\vec{q}) \cdot \underline{\vec{\delta}}_h \,\mathrm{d}\vec{x} + \int_T u \,\mathrm{div}\, \underline{\vec{\delta}}_h \,\mathrm{d}\vec{x} = \int_E u(\lambda_h - P_h^{\mathcal{E}} u)\,\mathrm{d}E\,,$$

which, together with the definitions of the projections P_h and $P_h^{\mathcal{E}}$, leads to

$$\int_T a^{-1}\vec{q} \cdot \underline{\vec{\delta}}_h \,\mathrm{d}\vec{x} - \int_T P_h u \,\mathrm{div}\, \underline{\vec{\delta}}_h \,\mathrm{d}\vec{x} + \int_E P_h^{\mathcal{E}} u(\lambda_h - P_h^{\mathcal{E}} u)\,\mathrm{d}E = \int_T \vec{g} \cdot \underline{\vec{\delta}}_h \,\mathrm{d}\vec{x} \tag{3.31}$$

Now, subtracting this from the first equation of (3.26) with test function

$$\vec{v}_h = \underline{\vec{\delta}}_h \text{ in } T \quad \text{and} \quad \vec{v}_h = \vec{0} \text{ in } \Omega \setminus T,$$

from (3.31) and using the scaling argument (3.39), we get

$$\begin{aligned}
\|\lambda_h - P_h^{\mathcal{E}} u\|_{L^2(E)}^2 &= \int_T a^{-1}(\vec{q} - \vec{q}_h) \cdot \underline{\vec{\delta}}_h \,\mathrm{d}\vec{x} - \int_T (P_h u - u_h) \mathrm{div}\, \underline{\vec{\delta}}_h \,\mathrm{d}\vec{x} \\
&\le \frac{1}{a_{\min}(\omega)} \|\vec{q} - \vec{q}_h\|_{L^2(T)} \|\underline{\vec{\delta}}_h\|_{L^2(T)} + \|u_h - P_h u\|_{L^2(T)} \|\mathrm{div}\, \underline{\vec{\delta}}_h\|_{L^2(T)} \\
&\lesssim \left(\frac{1}{a_{\min}(\omega)} h_T^{1/2} \|\vec{q} - \vec{q}_h\|_{L^2(T)} + h_T^{-1/2} \|u_h - P_h u\|_{L^2(T)} \right) \|\lambda_h - P_h^{\mathcal{E}} u\|_{L^2(E)}\,.
\end{aligned}$$

Dividing both sides of the estimate by $\|\lambda_h - P_h^{\mathcal{E}} u\|_{L^2(E)}$ gives the desired result. □

We now define an approximation $\widetilde{u}_{h,\omega} \in \mathcal{W}_{1;h}$, where $\mathcal{W}_{1;h} \subset L^2(D)$ denotes the Crouzeix-Raviart space of all (discontinuous) piecewise linear functions that have continuous averages across all element faces $E \in \mathcal{E}_I$ (cf. [**?**, § 1.2.6]). The construction is based on the following observation (cf. [**?**, Lemma 2.1], the proof of which extends in a straight forward way to $d = 3$).

Lemma 3.11 *Let $w \in L^2(\mathcal{E})$ and $T \in \mathcal{T}_h$. Then there exists a unique linear function $\chi_T = \chi_T(w)$ on T such that*

$$\int_E \chi_T \, dE = \int_E w \, dE \tag{3.32}$$

for all faces $E \subset \partial T$. Moreover,

$$\|\chi_T\|_{L^2(T)} \;\lesssim\; h_T^{1/2} \sum_{E \subset \partial T} \|w\|_{L^2(E)}\,. \tag{3.33}$$

The hidden constant depends only on the minimum angle of T.

We define the recovered pressure approximation $\widetilde{u}_{h,\omega} \in \mathcal{W}_{1;h}$ elementwise by

$$\widetilde{u}_{h,\omega}|_T := \chi_T\left(\lambda_{h,\omega}\right)\,, \quad \text{for all } \; T \in \mathcal{T}_h\,. \tag{3.34}$$

Lemma 3.11 ensures that $\widetilde{u}_{h,\omega} \in \mathcal{W}_{1;h}$ is well-defined and unique. The value of the multiplier $\lambda_{h,\omega}$ coincides with the value of $\widetilde{u}_{h,\omega}$ at the centre of each face $E \in \mathcal{E}$. Note that the recovered pressure approximation $\widetilde{u}_{h,\omega} \notin H_0^1(D)$ in general. It should not be confused with the standard continuous piecewise linear pressure approximation. The approximation error of $\widetilde{u}_{h,\omega}$ is bounded as follows.

Theorem 3.12 *Let u_ω be the pressure solution in* (2.1)-(2.2) *and let $\widetilde{u}_{h,\omega} \in \mathcal{W}_{1;h}$ be defined by* (3.34)*. Then, under the assumptions of Theorem 2.4 and for all $0 < s < \min(t, \frac{\pi}{\theta_{\max}}) < 1$,*

$$\|u_\omega - \widetilde{u}_{h,\omega}\|_{L^2(D)} \;\lesssim\; C_u(\omega)\, h^{2s} \tag{3.35}$$

where $C_u(\omega)$ is defined in (3.16)*. Moreover, $\|u - \widetilde{u}_h\|_{L^p(\Omega, L^2(D))} \lesssim h^{2s}$, for all $p < r$.*

Proof. We follow the proof of [**?**, Theorem 2.2] in the lowest order case for a fixed sample $\omega \in \Omega$, although we suppress the dependence on ω in the proof. We obtain the required estimate via the inequality

$$\|u - \widetilde{u}_h\|_{L^2(\Omega)} \;\leq\; \|u - \mathcal{I}_h^{\mathrm{CR}} u\|_{L^2(\Omega)} + \|\mathcal{I}_h^{\mathrm{CR}} u - \widetilde{u}_h\|_{L^2(\Omega)} \,, \tag{3.36}$$

where $\mathcal{I}_h^{\mathrm{CR}}$ is the (lowest order) Crouzeix-Raviart interpolation operator with respect to the mesh $\mathcal{T}_h$ (see, e.g., [**?**, § 1.4]).

To estimate the first term on the right-hand side of (3.36), first let $\widehat{T}$ denote the unit simplex in $\mathbb{R}^d$ and let $\widehat{\mathcal{I}}^{\mathrm{CR}}$ denote the Crouzeix-Raviart interpolation operator on $\widehat{T}$. Since the degrees of freedom of $\widehat{\mathcal{I}}^{\mathrm{CR}}\widehat{v}$ consist of integrals of $\widehat{v}$ on each of the faces of $\widehat{T}$, a simple application of Cauchy-Schwarz and the Trace Theorem yields the estimate: $\|\widehat{\mathcal{I}}^{\mathrm{CR}}\widehat{v}\|_{L^2(\widehat{T})} \lesssim \|\widehat{v}\|_{H^1(\widehat{T})}$ for all $\widehat{v} \in H^1(\widehat{T})$. Hence, for all affine polynomials $\widehat{p}$,

$$\|\widehat{u} - \widehat{\mathcal{I}}^{\mathrm{CR}}\widehat{u}\|_{L^2(\widehat{T})} \;=\; \|(\widehat{u} - \widehat{p}) - \widehat{\mathcal{I}}^{\mathrm{CR}}(\widehat{u} - \widehat{p})\|_{L^2(\widehat{T})} \;\lesssim\; \|\widehat{u} - \widehat{p}\|_{H^1(\widehat{T})},$$

and the Bramble Hilbert Lemma (e.g. [**?**, Thm 6.1]) then yields

$$\|\widehat{u} - \widehat{\mathcal{I}}^{\mathrm{CR}}\widehat{u}\|_{L^2(\widehat{T})} \;\lesssim\; \|\widehat{u}\|_{H^{1+s}(\widehat{T})}, \tag{3.37}$$

for all $s \in (0,1)$, where we take the norm on the right-hand side to be the Sobolev - Slobodeckij norm on $\widehat{T}$. Now, for all $v \in H^1(D)$ we have, by definition,

$$(\mathcal{I}_h^{\mathrm{CR}} v)(x) \;=\; (\widehat{\mathcal{I}}^{\mathrm{CR}}\widehat{v})(\kappa_T^{-1}(x)), \quad \text{for all} \quad x \in T \quad \text{and all} \quad T \in \mathcal{T}_h \,, \tag{3.38}$$

where κ_T denotes the affine map from $\widehat{T}$ to T and $\widehat{v} := v \circ \kappa_T$. Combining this with (3.37) and using the usual scaling argument yields

$$\|u - \mathcal{I}_h^{\mathrm{CR}} u\|_{L^2(T)} \;\lesssim\; h_T^{1+s} \|u\|_{H^{1+s}(T)}, \tag{3.39}$$

Then, since $s < 1$, squaring, summing and combining with Theorem 2.4, yields the required estimate for the first term on the right-hand side of (3.36). (Since h^{1+s} converges to zero more quickly than h^{2s} we do not need to compute the the exact asymptotic constant in the estimate (3.39).) We remark also that estimating the sum of the squares of the local fractional norms on T above by the square of the global fractional norm on D is in general is not sharp - see e.g. [**?**].

Turning to the second term in (3.36), recall that, for any $v \in H^1(D)$,

$$\int_E (v - \mathcal{I}_h^{\mathrm{CR}} v)\,\mathrm{d}E = 0\,, \qquad \text{for all faces} \;\; E \in \mathcal{E}. \tag{3.40}$$

Then it follows from the definition of $P_h^{\mathcal{E}}$, as well as from (3.34), (3.32) and (3.40) that

$$\int_E (\widetilde{u}_h - \mathcal{I}_h^{\mathrm{CR}} u)\,\mathrm{d}E = \int_E (\chi_T(\lambda_h) - u)\,\mathrm{d}E = \int_E (\lambda_h - P_h^{\mathcal{E}} u)\,\mathrm{d}E\,, \qquad \text{for all faces} \quad E \subset \partial T.$$

Thus, using Lemma 3.11 with $w = \lambda_h - P_h^{\mathcal{E}} u$ and $\chi_T = \widetilde{u}_h - \mathcal{I}_h^{\mathrm{CR}} u$ and combining it with the estimate in Lemma 3.10, we obtain for each element $T \in \mathcal{T}_h$,

$$\|\widetilde{u}_h - \mathcal{I}_h^{\mathrm{CR}} u\|_{L^2(T)} \lesssim h_T^{1/2} \sum_{E \subset \partial T} \|\lambda_h - P_h^{\mathcal{E}} u\|_{L^2(E)} \lesssim \frac{1}{a_{\min}(\omega)} h_T \|\vec{q} - \vec{q}_h\|_{L^2(T)} + \|u_h - P_h u\|_{L^2(T)}\,.$$

Now, squaring and summing over all $T \in \mathcal{T}_h$, the required estimate follows from Theorem 3.4 and Lemma 3.6.

The final estimate of the theorem follows from Assumptions A1–A3 and Hölder's inequality. □

Remark 3.13 Note that for diffusion coefficients with trajectories in $\mathcal{C}^t(\bar{D})$ with $t < 1/2$ (e.g. for the exponential covariance), for domains D with reentrant corners with $\theta_{\max} \approx 2\pi$, or for source terms f that are only in $L^2(D)$, the recovered pressure approximation $\widetilde{u}_h \in \mathcal{W}_{1;h}$ converges with the same rate as the piecewise constant approximation $u_h \in \mathcal{W}_h$ (see Theorem 3.7).

Remark 3.14 It is possible to prove the bound in (3.35) without using the intermediate result in Lemma 3.10. (See [**?**, § 7.4] for details.)

4 Application in the analysis of multilevel Monte Carlo methods

We apply the mixed FE error analysis carried out in Section 3 to the complexity analysis of *multilevel Monte Carlo methods*. Crucially, the FE error convergence rate determines the computational cost of multilevel Monte Carlo. It turns out that MLMC estimators are significantly more efficient than standard Monte Carlo for the groundwater flow problem. We use MLMC to estimate the expected value $\mathbb{E}[Q]$ of certain functionals $Q := \mathcal{G}(\vec{q}, u)$ of the solution to the lognormal diffusion problem (2.1)-(2.2). This approach is quite general since many important solution statistics, such as moments or failure probabilities, can be expressed in terms of expectations.

We fix a realisation $\omega \in \Omega$ and approximate the solution $(\vec{q}_\omega, u_\omega)$ to (2.1)-(2.2) by the associated mixed FE approximation $(\vec{q}_{h,\omega}, u_{h,\omega})$ which satisfies (3.1). This leads us to approximate the functional Q by $Q_h(\omega) := \mathcal{G}(\vec{q}_{h,\omega}, u_{h,\omega})$ evaluated on a FE mesh with mesh size h. A common method to estimate the expected value of $Q_h(\omega)$ is the *standard Monte Carlo estimator* for $\mathbb{E}[Q_h]$, defined as

$$\hat{Q}_{h,N}^{MC} := \frac{1}{N} \sum_{n=1}^{N} Q_h(\omega^{(n)}),$$

where $Q_h(\omega^{(n)})$ is the functional corresponding to sample $\omega^{(n)}$. Note that we compute N statistically independent samples in total.

Let $X : \Omega \to \mathbb{R}$ denote a random variable. In statistics it is common to quantify the accuracy of an estimator $\hat{X}$ to $\mathbb{E}[X]$ by the root mean square error (RMSE)

$$e(\hat{X})^2 := \mathbb{E}[\hat{X} - \mathbb{E}[X]]^2 \ .$$

The associated computational cost $C_\varepsilon(\hat{X})$ is characterised by the number of floating point operations required to achieve a RMSE $e(\hat{X}) \leq \varepsilon$.

The standard MC estimator is unbiased, $\mathbb{E}[\hat{Q}_{h,N}^{MC}] = \mathbb{E}[Q_h]$, and its variance is given by $\mathbb{V}[\hat{Q}_{h,N}^{MC}] = N^{-1}\mathbb{V}[Q_h]$. It is easy to see that these facts allow us to expand the mean square error (MSE) as

$$e(\hat{Q}_{h,N}^{MC})^2 = N^{-1}\mathbb{V}[Q_h] + (\mathbb{E}[Q_h - Q])^2 \ . \tag{4.1}$$

The first term above is the *sample error* and is mainly controlled by the sample size N. The second term, often referred to as *bias*, is determined solely by the FE approximation error. To ensure $e(\hat{Q}_{h,N}^{MC}) \leq \varepsilon$ it is sufficient that both terms in (4.1) are smaller than $\varepsilon^2/2$. For the sample error this can be achieved by choosing $N = O(\varepsilon^{-2})$. For the bias to be of order $|\mathbb{E}[Q_h - Q]| = O(\varepsilon)$, the FE mesh size has to be chosen sufficiently small. In the context of our groundwater flow problem this is a tall order. Since realisations of the diffusion coefficient are spatially rough and highly oscillatory we need a very fine FE mesh to obtain acceptable accuracies of the solution. At the same time, the number of samples is in general quite large due to the slow convergence of the Monte Carlo estimator.

The MLMC estimator overcomes this difficulty by estimating the expected value not only on a single FE mesh with fixed mesh size but on a *hierarchy* of increasingly finer FE meshes $\{\mathcal{T}_{h_\ell}\}_{\ell=0,\dots,L}$

with $\mathcal{T}_h := \mathcal{T}_{h_L}$ the finest mesh, and $h_\ell/h_{\ell-1} \leq c < 1$. Observe that by linearity of the expectation we may write

$$\mathbb{E}[Q_h] = \mathbb{E}[Q_{h_0}] + \sum_{\ell=1}^{L} \mathbb{E}[Q_{h_\ell} - Q_{h_{\ell-1}}] \ .$$

Let $Y_0 := Q_{h_0}$ and $Y_\ell := Q_{h_\ell} - Q_{h_{\ell-1}}$, $\ell = 1, \ldots, L$. The *multilevel Monte Carlo estimator* for $\mathbb{E}[Q_h]$ is then defined as

$$\hat{Q}_{h,N}^{ML} := \sum_{\ell=0}^{L} \hat{Y}_{\ell,N_\ell}^{MC} = \sum_{\ell=0}^{L} \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} Y_\ell(\omega^{(n)}) \ .$$

Since each correction Y_ℓ, $\ell = 0, \ldots, L$ is estimated independently from the others, the RMSE of the MLMC estimator reads

$$e(\hat{Q}_{h,N}^{ML})^2 = \sum_{\ell=0}^{L} N_\ell^{-1} \mathbb{V}[Y_\ell] + (\mathbb{E}[Q_h - Q])^2 \ .$$

Notably, the bias associated with the MLMC estimator has not changed compared to the standard MC estimator. However, the sample error can now be controlled and distributed over the entire hierarchy of FE meshes. Importantly, the variance of the differences $Y_\ell \to 0$ as $h_\ell \to 0$ and hence the number of samples N_ℓ required on the finer meshes is very small. Essentially, the MLMC estimator allows us to shift a large part of the computational effort to coarse, inexpensive grids and requires only a small number of expensive fine grid simulations while maintaining the same accuracy (in terms of the RMSE). A more detailed introduction to MLMC methods in the context of PDEs with random coefficients is presented in [**?**].

4.1 Multilevel Monte Carlo complexity

Let C_ℓ denote the cost to obtain one sample of Q_{h_ℓ}. The following result on the ε-cost of the MLMC estimator is taken from [**?**, Theorem 1].

Theorem 4.1 *Let $\alpha, \beta, \gamma, c_{M_1}, c_{M_2}, c_{M_3}$ be positive constants such that $\alpha \geq \frac{1}{2}\min\{\beta, \gamma\}$ and*

M1. $|\mathbb{E}[Q_{h_\ell} - Q]| \leq c_{M_1} h_\ell^\alpha$,

M2. $\mathbb{V}[Q_{h_\ell} - Q_{h_{\ell-1}}] \leq c_{M_2} h_\ell^\beta$,

M3. $C_\ell \leq c_{M_3} h_\ell^{-\gamma}$.

Then, for any $\varepsilon < e^{-1}$, there exist a level L and a sequence $\{N_\ell\}_{\ell=0}^L$, such that $e(\hat{Q}_{h,\{N_\ell\}}^{ML}) < \varepsilon$ and

$$C_\varepsilon(\hat{Q}_{h,\{N_\ell\}}^{ML}) \lesssim \begin{cases} \varepsilon^{-2}, & \text{if } \beta > \gamma, \\ \varepsilon^{-2}(\log \varepsilon)^2, & \text{if } \beta = \gamma, \\ \varepsilon^{-2-(\gamma-\beta)/\alpha}, & \text{if } \beta < \gamma \ . \end{cases}$$

The hidden constant depends on $c_{M_1}, c_{M_2}, c_{M_3}$.

In Theorem 4.1, M1-M2 describe assumptions on the spatial discretisation error and must be verified for each output quantity in conjunction with a particular spatial discretisation. We will prove these assumptions for the mixed FE discretisation of the lognormal diffusion problem and certain output quantities in Section 4.2 ahead. M3 is an assumption on the cost to obtain one

sample of the output Q_{h_ℓ}. In our problem this is the cost of obtaining one sample of the lognormal diffusion coefficient plus the cost to solve the associated discretised PDE problem.

Sampling the lognormal diffusion coefficient a can be done by computing approximate Karhunen-Loève eigenpairs of the underlying Gaussian random field $\log(a)$ (see e.g., [**?**, **?**]). Then, a certain number K_ℓ of leading eigenpairs is retained in the expansion on level ℓ and the truncated Karhunen-Loève expansion (KLE) of $\log(a)$ serves as approximation of the random field. The optimal choice of the truncation parameter K_ℓ that guarantees a negligible truncation error is problem-dependent. Typically we expect $K_\ell \gtrsim h_\ell^{-m}$, where $m = 1, 2$ (cf. [**?**, § 4.1]). Exact samples of the underlying Gaussian field can be obtained by computing a factorisation of the covariance matrix associated with the quadrature nodes on the FE mesh. A fast and efficient approach to do this is by circulant embedding (cf. [**?**]) which has at most log-linear complexity with respect to the number of quadrature points. Thus, assuming that the cost of the PDE solver is of optimal order, that is, it scales linearly with respect to the number of unknowns in the FE discretisation, then $\gamma \approx d$ (circulant embedding), or $\gamma \approx d + m$, $m = 1, 2$ (truncated KLE).

The three upper bounds in Theorem 4.1 correspond to three scenarios. Depending on the ratio of β and γ in Assumptions M2 and M3, the major part of the computational cost could be on the coarsest level ($\beta > \gamma$), spread evenly across all levels ($\beta = \gamma$) or on the finest level ($\beta < \gamma$).

In the context of realistic groundwater flow applications in two and three space dimensions the costs to obtain one sample grow rapidly with decreasing spatial resolution and thus we will almost always be in the last regime $\beta < \gamma$. If $\beta = 2\alpha$ (as is often the case, see Section 4.2 ahead), then, the total cost of the MLMC estimator is of order $\varepsilon^{-\gamma/\alpha}$ which is asymptotically the same as the cost to compute only *one* sample on the finest mesh to accuracy ε. The gains we can expect by using the MLMC estimator in place of the standard Monte Carlo estimator are thus significant and when $\beta = 2\alpha$ the MLMC estimator is asymptotically optimal.

4.2 Verifying the mixed FE error convergence rates

The performance analysis of the MLMC method as stated in Theorem 4.1 relies on bounds for $|\mathbb{E}[Q_{h_\ell} - Q]|$ and $\mathbb{V}[Q_{h_\ell} - Q_{h_{\ell-1}}]$ in terms of the characteristic mesh size h_ℓ on level ℓ. These bounds can be proved by bounding the spatial discretisation error. See [**?**, **?**] for an analysis of MLMC for (2.1)-(2.2) in the framework of standard FEs.

By using the appropriate error estimates derived in Section 3 we can now easily deduce the convergence rates α and β in Theorem 4.1 for various quantities of interest Q and the associated FE approximation Q_h. In Proposition 4.3 below, we use the following simple Lemma.

Lemma 4.2 *Let $\mathcal{B}$ denote a Banach space with norm $\|\cdot\|_{\mathcal{B}}$. Let $X, Y \in L^2(\Omega, \mathcal{B})$ be $\mathcal{B}$-valued random variables. Then we have*

$$|\mathbb{E}[\|X\|_{\mathcal{B}} - \|Y\|_{\mathcal{B}}]| \leq \|X - Y\|_{L^1(\Omega,\mathcal{B})}, \tag{4.2}$$

$$\mathbb{V}[\|X\|_{\mathcal{B}} - \|Y\|_{\mathcal{B}}] \leq \|X - Y\|^2_{L^2(\Omega,\mathcal{B})} . \tag{4.3}$$

Proof. It is easy to see that

$$|\mathbb{E}[\|X\|_{\mathcal{B}} - \|Y\|_{\mathcal{B}}]| \leq \mathbb{E}[|\|X\|_{\mathcal{B}} - \|Y\|_{\mathcal{B}}|] \leq \mathbb{E}[\|X - Y\|_{\mathcal{B}}] = \|X - Y\|_{L^1(\Omega,\mathcal{B})} ,$$

where, in the second step, we have used the reverse triangle inequality. To bound the variance of the difference we use $\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \leq \mathbb{E}[X^2]$, and, again, the reverse triangle inequality:

$$\mathbb{V}[\|X\|_{\mathcal{B}} - \|Y\|_{\mathcal{B}}] \leq \mathbb{E}[(\|X\|_{\mathcal{B}} - \|Y\|_{\mathcal{B}})^2] \lesssim \mathbb{E}[\|X - Y\|^2_{\mathcal{B}}] = \|X - Y\|^2_{L^2(\Omega,\mathcal{B})} .$$

□

Proposition 4.3 *Let $D \subset \mathbb{R}^2$ be a polygon with largest interior angle $\theta_{\max} \in (0, 2\pi)$. Let Assumptions A1-A3 hold for some $0 < t < 1$ and $r > 2$. Define $t^\star := \min(t, \frac{\pi}{\theta_{\max}})$. Then, assumptions M1-M2 hold with α and β as follows:*

Q	Q_h		Reference
$\|\vec{q}\|_{L^2(D)}$	$\|\vec{q}_h\|_{L^2(D)}$	$\alpha < t^\star,\ \beta < 2t^\star$	Corollary 3.5
$\|\vec{q}\|_{H(\mathrm{div})}$	$\|\vec{q}_h\|_{H(\mathrm{div})}$	$\alpha < t^\star,\ \beta < 2t^\star$	Corollary 3.5
$\|u\|_{L^2(D)}$	$\|u_h\|_{L^2(D)}$	$\alpha < \min(1, 2t^\star),\ \beta < \min(2, 4t^\star)$	Theorem 3.7
$\|u\|_{L^2(D)}$	$\|\widetilde{u}_h\|_{L^2(D)}$	$\alpha < 2t^\star,\ \beta < 4t^\star$	Theorem 3.12
$\mathcal{M}(\vec{q})$	$\mathcal{M}(\vec{q}_h)$	$\alpha < 2t^\star,\ \beta < 4t^\star$	Theorem 3.8

Here, $\mathcal{M} \in \mathcal{V}^\star$ is a linear functional of the Darcy velocity $\vec{q}$ in (2.3) that satisfies Assumption A4.

Proof. The first four cases follow by combining Lemma 4.2 with the corresponding error estimates in Section 3, choosing X, Y and $\mathcal{B}$ accordingly.

For linear velocity functionals $\mathcal{M}$ we obtain

$$|\mathbb{E}[\mathcal{M}(\vec{q} - \vec{q}_h)]| \le \mathbb{E}[|\mathcal{M}(\vec{q} - \vec{q}_h)|] = \|\mathcal{M}(\vec{q} - \vec{q}_h)\|_{L^1(\Omega)}$$

and

$$\mathbb{V}[\mathcal{M}(\vec{q} - \vec{q}_h)] \le \mathbb{E}[\mathcal{M}^2(\vec{q} - \vec{q}_h)] = \|\mathcal{M}(\vec{q} - \vec{q}_h)\|^2_{L^2(\Omega)} .$$

The assertion follows by combining these bounds with the error estimates in Theorem 3.8. □

Remark 4.4 Most of the quantities of interest in Proposition 4.3 are of limited physical relevance and are included primarily for their mathematical value. However, an important example of a linear functional $\mathcal{M}(\vec{q})$ of the Darcy velocity is given in (5.4) below. As stated in Section 2.2 and in Remark 3.9, it is possible to extend the results also to Fréchet-differentiable nonlinear functionals of pressure and velocity, to piecewise $C^t(\overline{D})$ coefficients and to $D \subset \mathbb{R}^3$ following [**?**, **?**].

5 Numerical experiments

As a representative example, we consider the 2D "flow cell" problem in mixed formulation

$$\begin{aligned} a^{-1}(\omega, \vec{x})\vec{q}(\omega, \vec{x}) + \nabla u(\omega, \vec{x}) &= 0, \\ \nabla \cdot \vec{q}(\omega, \vec{x}) &= 0, \qquad \text{in} \quad D = (0,1) \times (0,1). \end{aligned} \tag{5.1}$$

The horizontal boundaries are no-flow boundaries, that is, $\vec{\nu} \cdot \vec{q} = 0$. We have $u \equiv 1$ along the western (inflow) and $u \equiv 0$ along the eastern (outflow) boundary.

The diffusion coefficient $a(\omega, x)$ is a lognormal random field; $\log(a)$ is a mean-zero Gaussian random field with variance $\sigma^2 \equiv 1$ and a specific covariance function ρ. In our examples we will use the exponential covariance

$$\rho(r) = \rho_{\exp}(r) := \sigma^2 \ \exp(-r/\lambda), \tag{5.2}$$

where $r = \|\vec{x} - \vec{y}\|_2$ is the Euclidean distance of $\vec{x}, \vec{y} \in \mathbb{R}^d$, and $\lambda > 0$ denotes the correlation length. We will also consider the Matérn covariance function

$$\rho(r) = \rho_\nu(r) := \sigma^2 \ \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{r}{\widetilde{\lambda}}\right)^\nu K_\nu \left(\frac{r}{\widetilde{\lambda}}\right), \tag{5.3}$$

where K_ν is the modified Bessel function of second kind and order ν, and $\widetilde{\lambda} = \lambda/(2\sqrt{\nu})$ denotes the scaled correlation length. Using the asymptotics of $K_{0.5}$ it is possible to show that $\rho_{0.5}$ is actually the exponential covariance $\rho_{\exp}$ with correlation length $\widetilde{\lambda}$ instead of λ.

The physical discretisation of the weak formulation associated with (5.1) is done with lowest order Raviart-Thomas mixed finite elements for the Darcy velocity $\vec{q}$ and piecewise constant elements for the pressure u (see Section 3) on a uniform mesh of $n \times n$ squares, each divided into two triangles. We solve the resulting saddle point problems using the efficient divergence-free reduction technique for Raviart-Thomas finite elements introduced in [**?**, **?**]. The associated symmetric positive-definite linear system is solved with the sparse direct solver implemented in Matlab. The theoretical cost of this is $O(h^{-3})$, but in practice it is often faster. In addition, we compute the recovered pressure approximation $\widetilde{u}_h$ from Section 3.4 and the effective permeability

$$k_{\text{eff}}(\omega) = \frac{\int_D q_1(\omega, \vec{x})\, \mathrm{d}\vec{x}}{-\int_D \frac{\partial u}{\partial x_1}(\omega, \vec{x})\, \mathrm{d}\vec{x}}, \tag{5.4}$$

which simplifies to $\int_D q_1(\omega, \vec{x})\, \mathrm{d}\vec{x}$ for the flow cell problem (see [**?**]).

For the exponential covariance, we generate samples of $\log(a)$ at the vertices of the FE mesh using the circulant embedding technique (see, e.g. [**?**, section 5]). For the Matérn covariance we use a truncated Karhunen-Loève expansion (KLE) where only a certain number of the leading eigenpairs is retained. The eigenpairs are approximated by a spectral collocation method [**?**].

Remark 5.1 The analysis in Sections 2.2 and 3 has been performed for pure Dirichlet boundary conditions. The flow cell test problem features mixed Dirichlet/Neumann boundary conditions. This case is handled in [**?**, Remark 5.4 (b)] for the primal formulation of (5.1) and this analysis carries over to mixed formulations. Importantly, in the flow cell problem, no additional singularities are introduced at the points where the Dirichlet and Neumann boundary segments meet. To see this, we reflect the problem and the solution across the Neumann boundary and apply the regularity theory to the union of the original and reflected domain. Crucially, all angles of the new domain are less than $\pi/2$ and thus the full regularity of the pressure is maintained.

Remark 5.2 In hybridised mixed methods it is standard to compute the Lagrange multipliers first and subsequently recover the Darcy velocity and pressure by local post-processing. The piecewise linear pressure recovery is then also a cheap, local procedure. Similarly, in non-hybridised methods, the Lagrange multipliers and thus the piecewise linear pressure approximation $\widetilde{u}_{h,\omega}$ can be recovered cheaply by substituting the Darcy velocity $\vec{q}_{h,\omega} \in \mathcal{V}_h \subset RT_{-1}(\mathcal{T}_h)$ and the pressure $u_{h,\omega} \in \mathcal{W}_h$ into the first equation in (3.26) and solving for $\lambda_{h,\omega}$. Note that for the usual choices of basis functions in $RT_{-1}(\mathcal{T}_h)$, $\mathcal{W}_h$ and $M_{-1}(\mathcal{E}_I)$, this recovery is completely local.

5.1 Mixed FE error convergence rates

We investigate the convergence of FE approximations to (5.1) with respect to the characteristic mesh size $h_\ell = n_\ell^{-1}$, on a sequence of uniform meshes with $n_\ell = n_0 * 2^\ell$, $\ell = 0, \ldots, L$. We use a Monte Carlo method with $N = 2000$ samples for the convergence tests. The reference solutions $\vec{q}_\star$ and $u_\star$ are computed on a grid with $h_\star = 1/256$.

First, we consider the exponential covariance function $\rho = \rho_{\exp}$ in (5.2). In this case the trajectories of $\log(a)$ (and thus a) belong to $\mathcal{C}^t(\bar{D})$ almost surely for all $t < 1/2$ and Assumption A2 is satisfied for all $0 < t < 1/2$. Note that $\operatorname{div} \vec{q} \equiv 0$ and thus $\|\vec{q}\|_{H(\operatorname{div},D)} = \|\vec{q}\|_{L^2(D)}$. Consequently, our theory in Section 3 tells us to expect $\|\vec{q}_\star - \vec{q}_h\|_{L^2(\Omega, H(\operatorname{div},D))} = O(h^{1/2-\delta})$ and $\|u_\star - u_h\|_{L^2(\Omega,L^2(D))} = O(h)$, for all $\delta > 0$. We expect essentially the same convergence

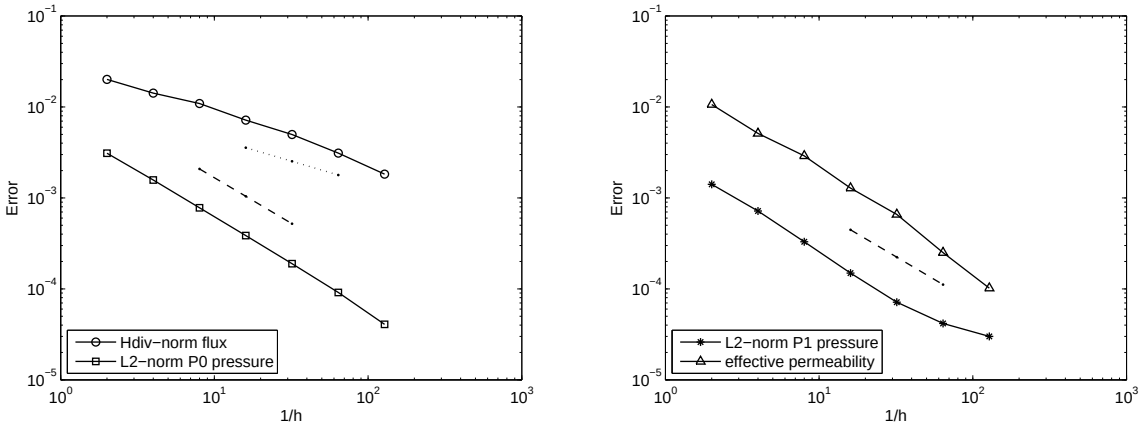

Figure 1: Mixed FE approximation errors with respect to the mesh size h_ℓ for $\rho = \rho_{\text{exp}}$ and $\lambda = 1$. The dotted line has slope $-1/2$; the dashed line has slope -1.

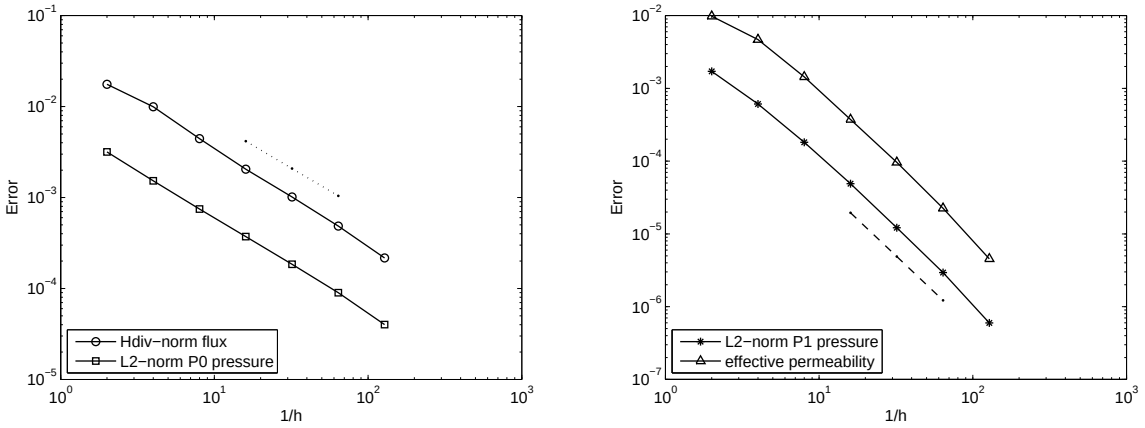

Figure 2: Mixed FE approximation errors with respect to the mesh size h_ℓ for $\rho = \rho_2$ and $\lambda = 0.5$. The dotted line has slope -1; the dashed line has slope -2.

$\|u_\star - \widetilde{u}_h\|_{L^2(\Omega, L^2(D))} = O(h^{1-\delta})$ for the recovered pressure approximation. For the effective permeability we also expect $|k_{\text{eff};\star} - k_{\text{eff};h}|_{L^2(\Omega)} = O(h^{1-\delta})$. The results in Figure 1 confirm our theory. We observe linear convergence for the standard and the recovered pressure approximation as well as the effective permeability. The velocity approximation is of order $O(h^{1/2})$.

Next we consider $\log(a)$ with Matérn covariance (5.3) with $\nu = 2$ and $\lambda = 0.5$. In this case it can be shown that the trajectories of $\log(a)$ belong to $\mathcal{C}^1(\bar{D})$ almost surely; hence Assumption A2 is satisfied for $t = 1$. We retain the leading 13 eigenpairs of the KLE which captures more than 95% of the variability of $\log(a)$. Due to the theory in Section 3 we expect linear convergence for both the velocity and standard pressure approximation and quadratic convergence for the recovered pressure approximation and the effective permeability. The results in Figure 2 confirm this. Note that these convergence rates are optimal for lowest-order Raviart Thomas mixed finite elements and can be achieved for all Matérn covariances with parameter $\nu > 1$.

5.2 Multilevel Monte Carlo simulation

We choose $\rho = \rho_{\text{exp}}$ and $\lambda = 0.1$ in our test problem. This time, we investigate the theoretical assumptions (M1) and (M2) in the MLMC complexity theorem for various quantities of interest derived from solutions to (5.1) using $N = 5000$ samples for the tests. The reference solution $(\vec{q}_\star, u_\star)$ along with $k_{\text{keff},\star}$ are again computed on a grid with $h_\star = 1/256$. The quantities of interest are the $H(\text{div})$-norm of the Darcy velocity and the effective permeability k_{eff} in (5.4).

Since the trajectories of a belong to $\mathcal{C}^t(\bar{D})$ almost surely for all $t < 1/2$, our theory in Section 4 tells us to expect $|\mathbb{E}[\|\vec{q}_h\|_{H(\text{div})} - \|\vec{q}_\star\|_{H(\text{div})}]| = O(h^{1/2})$ and $\mathbb{V}[\|\vec{q}_h\|_{H(\text{div})} - \|\vec{q}_{2h}\|_{H(\text{div})}] = O(h)$. The results in Figure 3 show that this result is not sharp. We observe twice the expected convergence rate. This is because $\|\vec{q}\|_{H(\text{div})} = (\int_D |\vec{q}|^2)^{1/2}$ here, which is a simple nonlinear functional of $\vec{q}$ that can be analysed by the same techniques as presented in [**?**] for standard piecewise linear FEs. For the effective permeability we expect $|\mathbb{E}[k_{\text{eff};h} - k_{\text{eff};\star}]| = O(h)$ and $\mathbb{V}[k_{\text{eff};h} - k_{\text{eff};2h}] = O(h^2)$. This time, the results in Figure 4 confirm this theory.

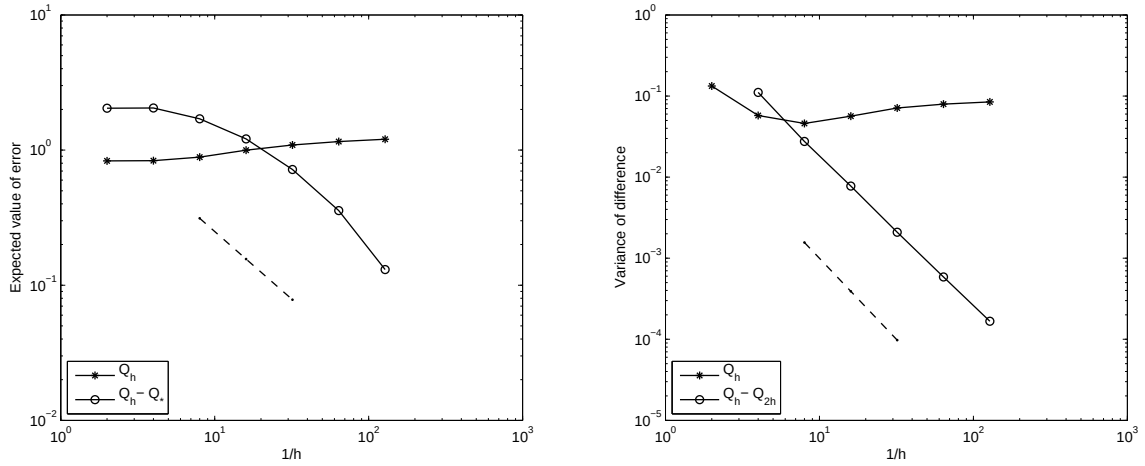

Figure 3: Plot of $|\mathbb{E}[\|\vec{q}_h\|_{H(\text{div})}]|$ and $|\mathbb{E}[\|\vec{q}_h\|_{H(\text{div})} - \|\vec{q}_\star\|_{H(\text{div})}]|$ (left), as well as $\mathbb{V}[\|\vec{q}_h\|_{H(\text{div})}]$ and $\mathbb{V}[\|\vec{q}_h\|_{H(\text{div})} - \|\vec{q}_{2h}\|_{H(\text{div})}]$ (right) with respect to the mesh size h for $\rho = \rho_{\text{exp}}$ and $\lambda = 0.1$. The dashed line has slope -1 (left) resp. -2 (right).

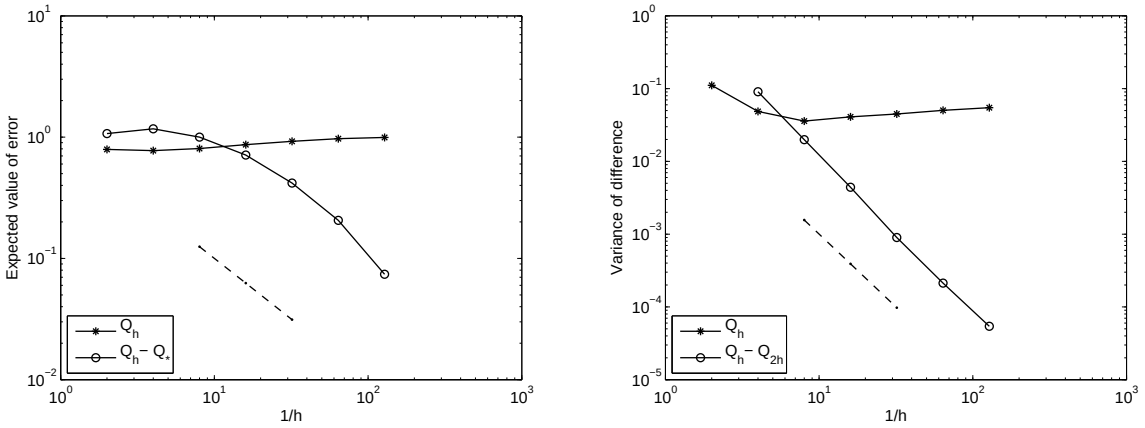

Figure 4: Plot of $|\mathbb{E}[k_{\text{eff};h}]|$ and $|\mathbb{E}[k_{\text{eff};h} - k_{\text{eff};\star}]|$ (left), as well as $\mathbb{V}[k_{\text{eff};h}]$ and $\mathbb{V}[k_{\text{eff};h} - k_{\text{eff};2h}]$ (right) with respect to the mesh size h for $\rho = \rho_{\text{exp}}$ and $\lambda = 0.1$. The dashed line has slope -1 (left) resp. -2 (right).

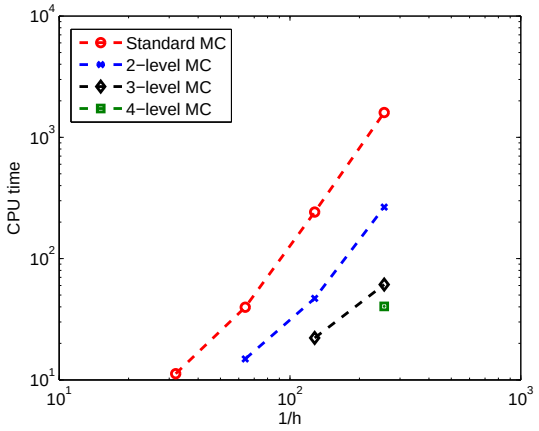

Figure 5: Plot of CPU time versus $1/h$ for a fixed tolerance of 10^{-4} for the sampling error. The quantity of interest is $k_{\text{eff};h}$. The correlation function is as in Figure 4.

Finally, in Figure 5 we compare the computational costs of the standard and multilevel MC estimator, respectively, for the estimation of the expected value of the effective permeability (5.4) where the sample error is fixed at 10^{-4} (i.e. $\varepsilon^2/2 = 10^{-4}$ and so $\varepsilon = \sqrt{2}10^{-2}$. The CPU timings (in seconds) shown in Figure 5 are calculated using Matlab 7.8 (with a single computational thread) on an 8 processor Linux machine with 14.6 GByte of RAM. As usual, the optimal number of samples N_ℓ in the MLMC estimator on each level is computed from the formula

$$N_\ell := \left\lceil 2\varepsilon^{-2}\sqrt{\widehat{s}_\ell^2 h_\ell}\sum_{\ell'=0}^{L}\sqrt{\widehat{s}_{\ell'}^2/h_{\ell'}}\right\rceil = 10^{-4}\left\lceil \sum_{\ell'=0}^{L}\sqrt{2^{\ell'-\ell}\,\widehat{s}_{\ell'}^2\widehat{s}_\ell^2}\right\rceil , \tag{5.5}$$

where $\widehat{s}_\ell^2$ is the sample variance on level ℓ, i.e. an estimate of $\mathbb{V}[Y_\ell]$ with an initial set of $\widetilde{N}_\ell < N_\ell$ samples (cf. [**?**, Section 5]).

The efficiency of MLMC as compared to standard MC is clearly demonstrated. On a mesh with $h = 1/256$ the 4-level MLMC estimator takes only 40 seconds, but the single level MC estimator takes 27 minutes. We observe that in Figure 4 the graphs of $\mathbb{V}[k_{\text{eff};h} - k_{\text{eff};2h}]$ and $\mathbb{V}[k_{\text{eff};h}]$ intersect at $h \approx \lambda = 0.1$. This means that the cost of the MLMC estimator on any coarser mesh will actually be bigger than the cost of the standard MC estimator on the same mesh. The coarsest mesh used in Figure 5 contains 32×32 elements.

This phenomenon was already observed in [**?**] and can be overcome by using smoother approximations of the permeability on coarser finite element meshes as suggested in [**?**, Section 4], where so called "level-dependent" MLMC estimators are studied. In [**?**] the log-permeability is approximated by a truncated KL expansion where a decreasing number of modes is included on the coarser meshes. A similar strategy has been suggested in the context of the related Brinkman problem in [**?**] under the assumption of a certain decay rate of the finite element error with respect to the number of KL modes. We do not study level-dependent MLMC estimators here since the focus of this work is on the mixed finite element error estimation.

5.3 Travel time calculations

Coming back to our motivating example in the introduction we now estimate statistics of the time it takes particles to travel from a location in the computational domain to its boundary. To do this we use a very simple particle tracking model. Having computed the Darcy velocity $\vec{q}_\omega$ via

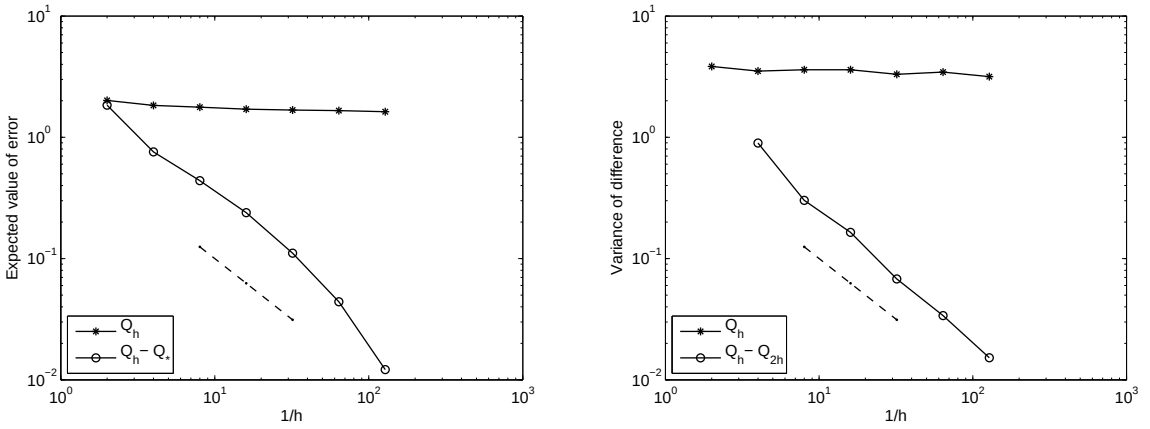

Figure 6: Plot of $|\mathbb{E}[\tau_h]|$ and $|\mathbb{E}[\tau_h - \tau_\star]|$ (left), as well as $\mathbb{V}[\tau_h]$ and $\mathbb{V}[\tau_h - \tau_{2h}]$ (right) with respect to the mesh size h for $\rho = \rho_{\exp}$ and $\lambda = 1$. The dashed line has slope -1.

(2.3) and neglecting molecular dispersion, the particle path satisfies the initial value problem

$$\frac{d\vec{x}_\omega}{dt} = \vec{q}_\omega, \quad \vec{x}_\omega(0) = \vec{x}_0 \ , \tag{5.6}$$

where $\vec{x}_0 \in D$ denotes the starting point. The travel time $\tau_\omega \in [0, \infty)$ is the time when the particle hits the boundary, i.e. when $\vec{x}_\omega(\tau_\omega) \in \partial D$ for the first time.

Due to the Picard-Lindelöf Theorem, a sufficient condition for problem (5.6) to have a unique solution is that $\vec{q}_\omega$ is Lipschitz continuous on D as a function of $\vec{x}$. Assuming $u_\omega \in C^{1+t}(\bar{D})$ it follows that $\nabla u_\omega \in C^t(\bar{D})^d$. Combining this with Assumption A2, that is, $a_\omega \in C^t(\bar{D})$, for some $0 < t \leq 1$, we see that the Darcy velocity $\vec{q}_\omega = -a_\omega \nabla u_\omega \in C^t(\bar{D})^d$ is Hölder continuous on $\bar{D}$ with coefficient t. Thus, the Darcy velocity is Lipschitz only if $t = 1$. The required regularity result for the pressure $u_\omega \in C^{1+t}(\bar{D})$ can be proved under Assumptions A1-A2 together with slightly stronger regularity conditions on the source terms and on the boundary data, see [**?**, Sect. 2.6]. For the exponential covariance where $t < 1/2$, the Darcy velocity $\vec{q}_\omega$ is not Lipschitz and we cannot use the Picard-Lindelöf Theorem. In this case the existence of the exact particle trajectories is unclear.

To discretise problem (5.6) we replace $\vec{q}_\omega$ by its FE approximation $\vec{q}_{h,\omega}$ in (3.1) and obtain

$$\frac{d\vec{x}_{\omega,h}}{dt} = \vec{q}_{h,\omega}, \quad \vec{x}_{h,\omega}(0) = \vec{x}_0 \ . \tag{5.7}$$

Again, the Picard-Lindelöf Theorem tells us that problem (5.7) has a unique solution which can be computed element by element over the triangulation $\mathcal{T}_h$ of D. By following the particle through the domain D and summing up the travel times in each element we obtain the FE approximation $\tau_{h,\omega}$ to the actual travel time τ_ω for each realisation a_ω of the random permeability.

We use the 2D flow cell problem with exponential covariance (5.2) and $\lambda = 1$. The particles are released at $\vec{x}_0 = [0, 0.5]^\top$. The reference travel time $\tau_\star$ is computed on a grid with $n = 256$ elements in each spatial direction. The results are depicted in Figure 6 where we use again $N = 5000$ samples on each level. Interestingly, we observe linear convergence in both plots suggesting that $\alpha = \beta$ in Theorem 4.1 in this case, in contrast to all the previous examples where $\beta = 2\alpha$. This is due to a difference in weak convergence (in mean) needed in (M1) versus strong convergence (pathwise) needed in (M2) for the travel time. We remark that the same behaviour is observed in stopping time calculations associated with Itô stochastic differential equations.

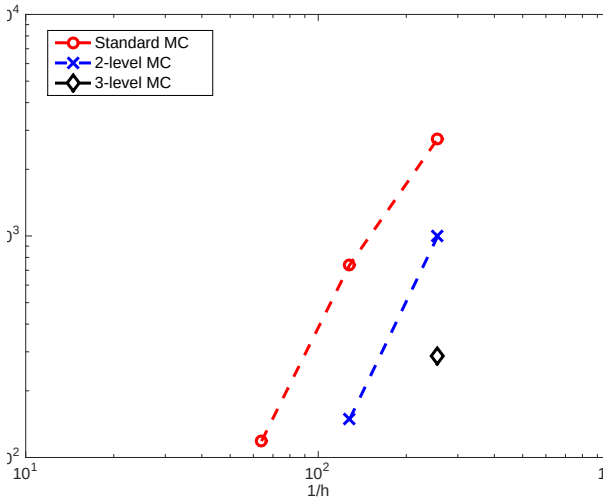

Figure 7: Plot of CPU time versus $1/h$ for a fixed tolerance 2.5×10^{-3} of the sampling error. The quantity of interest is $\mathbb{E}[\tau_h]$. This plot is for $\rho = \rho_{\text{exp}}$ and $\lambda = 1$.

Finally, we compare the computational costs of the standard and multilevel MC estimator, respectively, for the estimation of the expected value of the travel time τ where the sample error is fixed at 2.5×10^{-3} (i.e. $\varepsilon = 5\sqrt{2} \times 10^{-2}$). The CPU timings (in seconds) shown in Figure 7 are calculated using Matlab 7.13 on a quadcore Linux machine with a 3.30 GHz processor and 4 GByte of RAM with the optimal numbers of samples N_ℓ in (5.5). The efficiency of MLMC as compared to standard MC is clearly demonstrated. On a mesh with $h = 1/256$ the 3-level MLMC estimator takes only 5 minutes, but the single level MC estimator takes 46 minutes.

6 Conclusions

We studied a single phase flow problem in a random porous medium described by correlated lognormal distributions. Realisations of the permeability are not uniformly bounded away from zero and infinity and they are in general only Hölder-continuous with exponent $0 < t \leq 1$. We presented a mixed formulation of this problem and established the regularity of the Darcy velocity and the pressure. Using lowest order Raviart-Thomas mixed elements we proved finite element error bounds for the Darcy velocity and the pressure, as well as a recovered pressure approximation. We showed that piecewise linear pressure recovery is only of practical interest for conductivity models with $t > 1/2$. Moreover, we proved error bounds for a class of linear functionals of the velocity. This enabled us to bound the finite element approximation error for the effective permeability of a 2D flow cell.

Using the finite element error bounds we also proved convergence of a Multilevel Monte Carlo

algorithm that is used to estimate the statistics of output quantities of interest associated with the flow problem, for example, the $H(\mathrm{div})$-norm of the velocity or the effective permeability. In addition, we estimated the expected value of particle travel times to the boundary of the computational domain. As for all other output quantities in this work we observe that the MLMC estimator outperforms standard Monte Carlo. However, the gains are less substantial and there is still potential for improvement using ideas developed in the context of stochastic ordinary differential equations, e.g. in [**?**], which merits further investigation.

Acknowledgements

We thank Daniele Boffi and Rodolfo Rodriguez for helpful comments. We also thank the anonymous referee for their suggestions that have led to an improvement of the proof of Theorem 3.12. This research was supported by the Engineering and Physical Sciences Research Council, UK, under grant EP/H051503/1.